THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Une généralisation, au cas non différentiable, de la matrice hessienne d'une fonction convexe : le différentiel second

LEVECQ, Sophie

Award date:
1991

Awarding institution:
Universite de Namur

FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX
NAMUR
FACULTE DES SCIENCES

UNE GENERALISATION, AU CAS NON DIFFERENTIABLE, DE LA MATRICE HESSIENNE D'UNE FONCTION CONVEXE : LE DIFFERENTIEL SECOND.

Promoteur : J.-J. STRODIOT

Sophie LEVECQ

Année académique : 1990-1991

Je tiens à remercier Monsieur le Professeur J.-J. STRODIOT pour son aide précieuse, sa disponibilité et son soutien constant tout au long de l'élaboration de ce travail.

J'adresse également toute ma reconnaissance à D. GOELEVEN pour sa collaboration.

Enfin, je voudrais remercier ma famille et mes amis pour leurs encouragements, leur attention et leur patience.

CHAPITRE 0

MOTIVATION

Considérons le problème d'optimisation suivant

$$(P) = \begin{cases} \min f(x) \\ \text{sous } x \in I\!R^n \end{cases}$$

où f représente une fonction convexe deux fois différentiable, définie sur $I\!R^n$ et à valeurs réelles. La méthode de Newton permet de le résoudre en remplaçant la fonction par son développement en série de Taylor, à l'ordre deux, au voisinage d'un point x_0 de $I\!R^n$. De cette façon, le problème se réécrit

$$\begin{cases} \min & f(x_0) + \nabla f(x_0)^T(x - x_0) + \frac{1}{2}(x - x_0)^T \nabla^2 f(x_0)(x - x_0) \\ \text{sous} & x \in I\!R^n \end{cases}$$

ou de manière équivalente

$$\begin{cases} \min & \nabla f(x_0)^T d + \frac{1}{2} d^T \nabla^2 f(x_0) d \\ \text{sous} & d \in I\!R^n \end{cases}$$

Notant sa solution d_0, nous déterminons le point x_1 de l'itération suivante au moyen de la relation $x_1 = x_0 + d_0$. Si la condition d'optimalité, soit $\nabla f(x_1) = 0$, n'est pas satisfaite, on calculera un second point x_2 suivant la même procédure, jusqu'à obtenir finalement, une suite (x_k) répondant au critère de convergence locale. La vitesse de convergence est quadratique et en celà, cette méthode est performante, donc désirable. Cependant, celle-ci ne s'applique qu'aux fonctions deux fois différentiables et est, en ce sens, assez restrictive.

Pour généraliser la méthode de Newton au cas non différentiable, nous pouvons envisager plusieurs possibilités :

- remplacer $\nabla^2 f(x_0)$ par un ensemble de matrices appelé matrice hessienne généralisée. Cette approche est détaillée dans [9].

- trouver un ensemble, noté $\partial^2 f(x_0)$, jouant un rôle similaire à celui de la dérivée du premier ordre généralisée, $\partial f(x_0)$.

A cet effet, nous introduisons la notion de dérivée directionnelle du second ordre de f au point x_0, soit $f''(x_0; d)$, comme substitut du terme $d^T \nabla^2 f(x_0) d$ dans le développement de Taylor de f au point x_0. Cette fonction de d n'étant en général ni convexe, ni semi-continue inférieurement, ne pourra s'exprimer comme fonction d'appui d'un certain ensemble K convexe, fermé, non vide. Dès lors, nous définirons pour tout x_0^* dans $\partial f(x_0)$, la dérivée directionnelle du second ordre de f en (x_0, x_0^*) dont la racine carrée satisfera l'équation

$$cl\sqrt{f''(x_0, x_0^*; .)}(d) = \delta^*(\partial^2 f(x_0, x_0^*) \mid d)$$

c'est-à-dire constituera, à l'opération de fermeture près, la fonction d'appui d'un ensemble K qu'Hiriart-Urruty appellera différentiel second de f en (x_0, x_0^*) et notera $\partial^2 f(x_0, x_0^*)$. Considérant alors l'intersection de tous les $\partial^2 f(x_0, x_0^*)$ pour $x_0^* \in \partial^2 f(x_0)$, nous générerons le différentiel second de f en x_0, soit $\partial^2 f(x_0)$. Il s'agira d'un ensemble convexe compact, contenant l'origine et représentant, dans le cas deux fois différentiable, l'ellipsoïde de $\mathbb{R}^n$, associé à la matrice hessienne $\nabla^2 f(x_0)$.

Cette définition du différentiel second de f au point x_0 pouvant, dans certains cas, s'avérer difficile à appliquer effectivement, nous introduirons des règles de calcul permettant de dériver l'expression du différentiel second d'une fonction f, construite à partir d'autres fonctions f_i dont les $\partial^2 f_i(x_0)$ sont plus faciles à calculer.

Nous présenterons ensuite les concepts de dérivée seconde intérieure et extérieure, obtenues en associant à $\partial^2 f(x_0)$ deux matrices symétriques semi-définies positives $\underline{\nabla}^2 f(x_0)$ et $\overline{\nabla}^2 f(x_0)$. Celles-ci correspondent, d'une part, à l'unique ellipsoïde de volume maximal centré en l'origine et inclus dans $\partial^2 f(x_0)$ et, d'autre part, à l'unique ellipsoïde de volume minimal centré en 0 et contenant $\partial^2 f(x_0)$. Dans le cas deux fois différentiable, ces matrices ne sont autres que la matrice hessienne de f au point x_0.

Finalement, nous procéderons à une comparaison entre le différentiel second $\partial^2 f(x_0)$ et la matrice hessienne généralisée au sens de Clarke, soit $\partial_c^2 f(x_0)$, où f représente ici une fonction convexe, différentiable, dont l'application gradient est localement lipschitzienne.

Notre approche s'effectuera dans un cadre relativement simple puisque nous ne considérerons que des fonctions convexes définies sur $\mathbb{R}^n$ et valeurs dans $\mathbb{R}$. Une généralisation des résultats se trouve développée dans [3]. On y considère un espace vectoriel général X et des fonctions convexes à valeurs dans $\mathbb{R} \cup \{+\infty\}$. Les idées de base restent cependant les mêmes.

Notre texte est organisé comme suit.

Chapitre I : Dérivées directionnelles du second ordre

1. Quelques rappels concernant le premier ordre

2. Le second ordre

- **2.1** Définition de $f''(x_0; d)$
- **2.2** Propriétés
- **2.3** Développement de Taylor et méthode de Newton
- **2.4** Dérivée directionnelle du second ordre et fonction d'appui
  - **2.4.1** La notion de fonction d'appui
  - **2.4.2** Propriétés de la forme quadratique $d \to d^T \nabla^2 f(x_0) d$ dans le cas deux fois différentiable
  - **2.4.3** Dérivée directionnelle du second ordre et fonction d'appui
    - **A)** Comment convexifier $\sqrt{f''(x_0; d)}$?
    - **B)** Comment rendre $\sqrt{f''(x_0, x_0^*; d)}$ semi-continue inférieurement ?

Chapitre II : Différentiels seconds

1. Différentiel second de f en (x_0, x_0^*)

2. Différentiel second de f en x_0

3. Résultats complémentaires et méthode de Newton

- **3.1** Différentiels seconds de deux fonctions comparables dans un voisinage de x
- **3.2** Développements du second ordre

CHAPITRE I

DERIVEES DIRECTIONNELLES

DU SECOND ORDRE

1 Quelques rappels concernant le premier ordre

Considérons une fonction f convexe définie sur $I\!R^n$ et à valeurs dans $I\!R$ ainsi qu'un point $x_0 \in I\!R^n$.
Rappelons qu'on appelle dérivée directionnelle de f au point x_0 dans la direction $d \in I\!R^n$, le nombre

$$f'(x_0; d) = \lim_{\lambda \to 0^+} \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}$$

Puisque la fonction est convexe, le quotient

$$\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}$$

est une fonction croissante en λ et nous pouvons écrire

$$f'(x_0; d) = \inf_{\lambda > 0} \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}$$

Lorsque f est différentiable en x_0, elle peut encore s'exprimer comme

$$f'(x_0; d) = \nabla f(x_0)^T d$$

où $\nabla f(x_0)$ désigne le gradient de f au point x_0. Par contre, lorsque f n'est plus différentiable en x_0, le vecteur gradient n'existe plus et est remplacé par l'ensemble

$$\partial f(x_0) = \{x^* \in I\!R^n \mid f(x) \geq f(x_0) + < x^*, x - x_0 > \quad \forall\, x \in I\!R^n\}$$

appelé le sous-différentiel de f en x_0.
En dimension 1, il représente l'ensemble des pentes des minorantes affines au graphe de f, exactes en x_0.

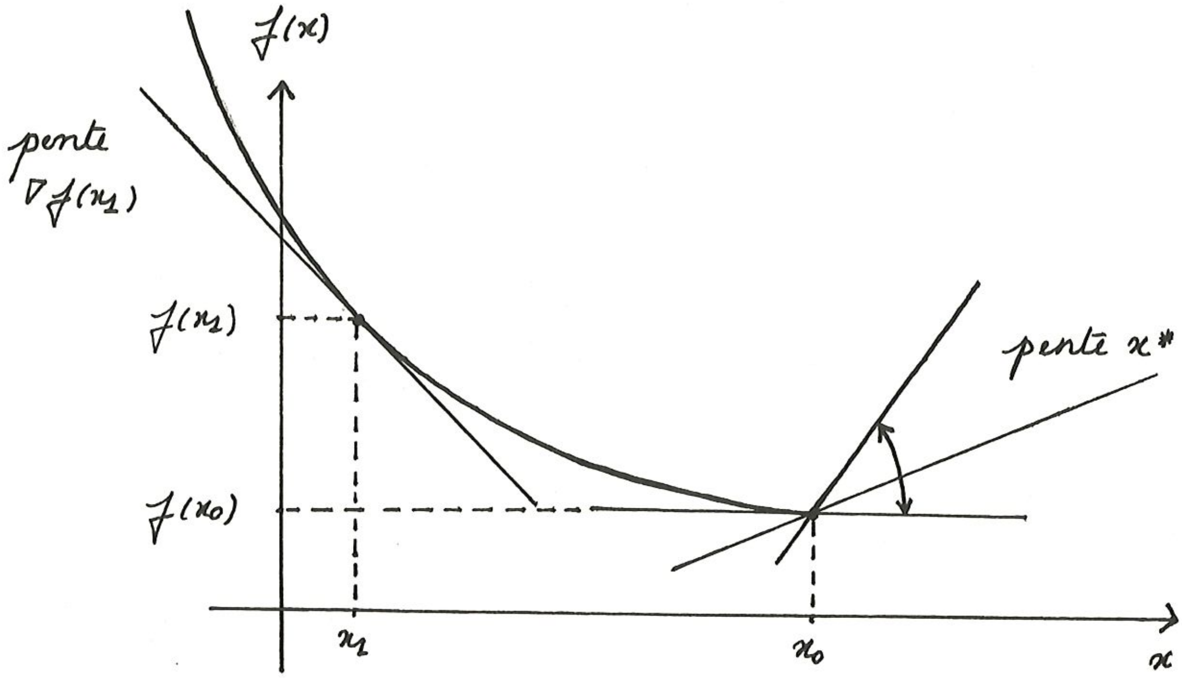


Figure 1

On trouvera en annexe quelques propriétés fondamentales des dérivées directionnelles et des sous-différentiels.

2 Le second ordre

Lorsque la fonction convexe $f : \mathbb{R}^n \to \mathbb{R}$ est deux fois différentiable en x_0, on peut écrire son développement de Taylor à l'ordre 2 autour de x_0 :

$\forall\, d \in \mathbb{R}^n,\ \forall\, \lambda > 0$

$$f(x_0 + \lambda d) = f(x_0) + \lambda \nabla f(x_0)^T d + \frac{1}{2}\lambda^2 d^T \nabla^2 f(x_0) d + o(\lambda^2)$$

où $\nabla^2 f(x_0)$ désigne la matrice hessienne de f en x_0.
Ce développement peut encore s'écrire

$$d^T \nabla^2 f(x_0) d = \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - \nabla f(x_0)^T d \right]$$

Lorsque f n'est pas deux fois différentiable, nous ne pouvons plus considérer la matrice hessienne de la fonction et devons chercher une solution de remplacement dans le développement ci-dessus. Plusieurs possiblités ont été soulevées dans le chapitre d'introduction et nous avons opté pour le remplacement direct du terme $d \rightarrow d^T \nabla^2 f(x_0) d$ par la dérivée directionnelle du second ordre de f en x_0 dans la direction d.

2.1 Définition

Nous appelons *dérivée directionnelle du second ordre* de f en x_0 dans la direction d, l'une quelconque des deux limites ci-dessous :

$$\lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - f'(x_0; d) \right] \tag{1}$$

ou

$$\lim_{\lambda \to 0^+} \frac{1}{\lambda} [f'(x_0 + \lambda d; d) - f'(x_0; d)] \tag{2}$$

Un résultat de Jessen (cfr [4]) nous permet en effet d'affirmer que la limite lorsque $\lambda \to 0^+$ de la quantité (1) existe si et seulement si elle existe pour (2). Cette limite commune sera notée $f''(x_0; d)$. Nous constatons que dans le cas deux fois différentiable, nous retrouvons la forme quadratique $d^T \nabla^2 f(x_0) d$ puisque la dérivée directionnelle du premier ordre $f'(x_0; d) = \nabla f(x_0)^T d$.

Nous supposerons dorénavant que $f''(x_0; d)$ existe et est finie pour tout x_0 et tout d. (Ces restrictions pourraient être levées en considérant la

$$\lim_{\lambda \to 0^+} \sup \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - f'(x_0; d) \right]$$

dans $I\!R^+ \cup \{+\infty\}$. Voir [3]).

Nous examinons à présent les caractéristiques principales de $f''(x_0; d)$.

2.2 Propriétés de $\mathbf{f''(x_0; d)}$

Proposition 2.2.1 :

$$f''(x_0; 0) = 0 \text{ et } f''(x_0; d) \geq 0 \ \forall \ d$$

Preuve :

La première assertion se vérifie immédiatement. Concernant la seconde, nous savons que

$$f''(x_0; d) = \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - \lim_{\mu \to 0^+} \frac{f(x_0 + \mu d) - f(x_0)}{\mu} \right]$$

et en utilisant l'hypothèse de convexité de la fonction f considérée, nous obtenons

$$f''(x_0; d) = \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - \inf_{\mu > 0} \frac{f(x_0 + \mu d) - f(x_0)}{\mu} \right]$$

où l'expression

$$\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - \inf_{\mu > 0} \frac{f(x_0 + \mu d) - f(x_0)}{\mu}$$

est positive. ∎

Proposition 2.2.2 :

$f''(x_0; \cdot)$ est positivement homogène de degré 2
c'est-à-dire $\forall \ \alpha \geq 0 \ f''(x_0; \alpha d) = \alpha^2 \ f''(x_0; d)$

Preuve :

Par définition,

$$f''(x_0; \alpha d) = \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda \alpha d) - f(x_0)}{\lambda} - \lim_{\lambda' \to 0^+} \frac{f(x_0 + \lambda' \alpha d) - f(x_0)}{\lambda'} \right]$$

et en considérant les changements de variable $\mu = \alpha\lambda$ et $\mu' = \alpha\lambda'$, nous établissons que

$$
\begin{aligned}
f''(x_0;\alpha d) &= \lim_{\mu\to 0^+} \frac{2\alpha}{\mu}\left[\frac{\alpha f(x_0+\mu d)-f(x_0)}{\mu}\right.\\
&\qquad\qquad \left. - \lim_{\mu'\to 0^+}\frac{\alpha f(x_0+\mu' d)-f(x_0)}{\mu'}\right]\\
&= \lim_{\mu\to 0^+} \frac{2\alpha}{\mu}\left[\frac{\alpha f(x_0+\mu d)-f(x_0)}{\mu} - \alpha f'(x_0;d)\right]\\
&= \alpha^2 f''(x_0;d)
\end{aligned}
$$

■

Nous pouvons cependant regretter qu'en général $f''(x_0;\cdot)$ ne soit ni convexe, ni semi-continue inférieurement comme nous le montre l'exemple suivant.

<u>**Exemple 1 :**</u> Considérons la fonction $f : I\!R^2 \to I\!R$ définie par

$$f(x) = \max\{0, x_1^2 + x_2^2 + 2x_1\} \quad \forall\, x = (x_1, x_2)$$

On a que

$$x_1^2 + x_2^2 + 2x_1 \leq 0$$

est équivalent à

$$(x_1+1)^2 + x_2^2 \leq 1$$

Dès lors, $f(x)$ sera nulle sur le disque D de centre $(-1,0)$ et de rayon 1, et vaudra

$$x_1^2 + x_2^2 + 2x_1$$

ailleurs.

Prenons $x_0 = (0,0)$ et $d = (d_1, d_2)$ et envisageons deux cas :

a) $d_1 < 0$:

On a ainsi $f(x_0 + \lambda d) = f(\lambda d_1, \lambda d_2)$ où $(\lambda d_1, \lambda d_2)$ appartient au disque D pour λ suffisamment petit.

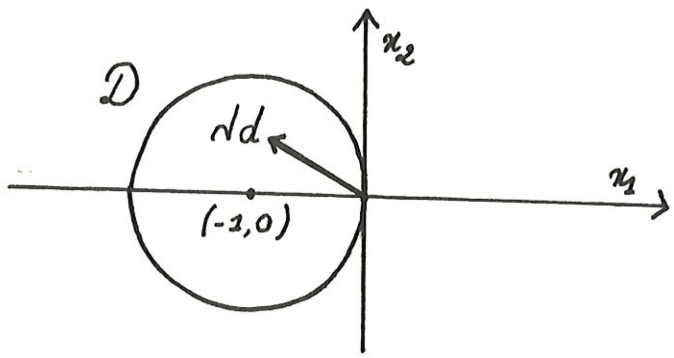


Figure 2

Autrement dit, $f(x_0 + \lambda d) = 0$ pour λ assez petit et, puisque

$$f(x_0) = f(0) = 0$$

nous obtenons

$$\begin{aligned} f''(x_0; d) &= \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} \right. \\ &\qquad \left. - \lim_{\mu \to 0^+} \frac{f(x_0 + \mu d) - f(x_0)}{\mu} \right] \\ &= \lim_{\lambda \to 0^+} \frac{2}{\lambda} [0 - 0] = 0 \end{aligned}$$

b) $d_1 \geq 0$:

Dans ce cas, $f(x_0 + \lambda d) = f(\lambda d_1, \lambda d_2)$ où $(\lambda d_1, \lambda d_2)$ ne se trouve pas dans le disque D. Ainsi,

$$f(x_0 + \lambda d) = \lambda^2 d_1^2 + \lambda^2 d_2^2 + 2\lambda d_1$$

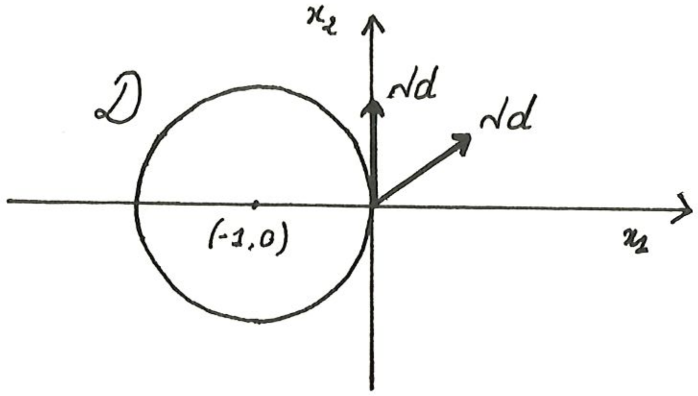


Figure 3

Par conséquent,

$$\begin{aligned} f''(x_0;d) &= \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{\lambda^2 d_1^2 + \lambda^2 d_2^2 + 2\lambda d_1 - 0}{\lambda} \right. \\ & \qquad \left. - \lim_{\mu \to 0^+} \frac{\mu^2 d_1^2 + \mu^2 d_2^2 + 2\mu d_1 - 0}{\mu} \right] \\ &= \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{\lambda^2 d_1^2 + \lambda^2 d_2^2 + 2\lambda d_1 - 2d_1}{\lambda} \right] \\ &= 2d_1^2 + 2d_2^2 \end{aligned}$$

Nous aurions encore pu obtenir ce résultat en constatant que

$$f_2(x) \overset{\text{not}}{=} x_1^2 + x_2^2 + 2x_1$$

constitue une fonction deux fois différentiable. Ainsi,

$$\begin{aligned} \forall\ d_1 \geq 0 \ \ f''(x_0;d) &= d^T \nabla^2 f_2(x_0) d \\ &= (d_1\ d_2) \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \\ &= 2(d_1^2 + d_2^2) \end{aligned}$$

Finalement, nous obtenons pour la dérivée directionnelle du second ordre

$$f''(x_0;d) = \begin{cases} 0 & \text{si } d_1 < 0 \\ 2(d_1^2 + d_2^2) & \text{si } d_1 \geq 0 \end{cases}$$

qui ne constitue évidemment pas une fonction convexe en d [cfr A.I.8].
Vérifions encore que $f''(x_0;\cdot)$ n'est pas semi-continue inférieurement aux points

$$\bar{d} = (0, \bar{d}_2) \text{ avec } \bar{d}_2 \neq 0$$

Il nous faut donc prouver [cfr A.II.1] que la propriété suivante est mise en défaut

$$\forall\, \lambda < f''(x_0; (0, \bar{d}_2)) \;\; \exists\, \delta > 0 \text{ t.q. } ||d - \bar{d}|| < \delta \Rightarrow \lambda < f_0''(x_0; d)$$

ce qui peut se réécrire

$$\forall\, \lambda < 2\bar{d}_2^2 \;\; \exists\, \delta > 0 \text{ t.q.}$$

$$d \in B(\bar{d}, \delta) \Rightarrow \lambda < \begin{cases} 0 & \text{si } d_1 < 0 \\ 2(d_1^2 + d_2^2) & \text{si } d_1 \geq 0 \end{cases}$$

où $B(\bar{d}, \delta)$ représente la boule ouverte de centre $\bar{d}$ et de rayon δ.
Représentons géométriquement $f''(x_0; d)$:

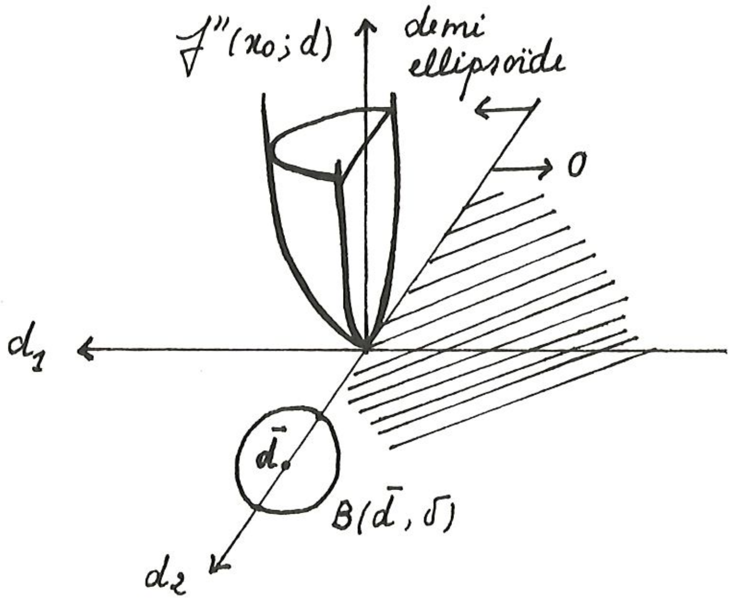


Figure 4

Nous constatons alors qu'il nous suffit de prendre $d = (d_1, d_2)$ avec $d_1 < 0$. Ainsi,

$$\forall\, \lambda < 2\bar{d}_2^2 \;\; \exists\, \delta > 0 \text{ t.q. } d \in B(\bar{d}, \delta) \not\Rightarrow \lambda < 0$$

■

Nous pouvons, à présent, nous pencher sur l'étude de deux cas particuliers pour lesquels $f''(x_0;\cdot)$ est convexe.

Proposition 2.2.3 : **[Le cas polyédral]**

Soit la fonction polyédrale définie par

$$f(x) = \max_{1\leq i\leq p} \{a_i x + b_i\}$$

Nous avons que $f''(x_0; d) = 0 \ \ \forall\, d \in I\!R^n$

Preuve :

En vertu de (1), il nous suffit de démontrer que

$$\exists\, \lambda_0 > 0 \text{ t.q. } \forall\, \lambda \leq \lambda_0 \ \ f'(x_0; d) = \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}$$

Considérons alors la fonction φ définie sur $I\!R^+$ et à valeurs dans $I\!R$ t.q.

$$\varphi(\lambda) = f(x_0 + \lambda d)$$

et montrons que :

$$\exists \lambda_0 > 0 \text{ t.q. } \forall\, \lambda \leq \lambda_0 \ \ \varphi'_+(0) = \frac{\varphi(\lambda) - \varphi(0)}{\lambda}$$

où $\varphi'_+(0)$ désigne la dérivée à droite de φ au point 0. La thèse s'en suivra.
Puisque φ est polyédrale, elle possède un nombre fini de points de non différentiabilité ou, d'une manière équivalente, son épigraphe contient un nombre fini de points extrémaux.

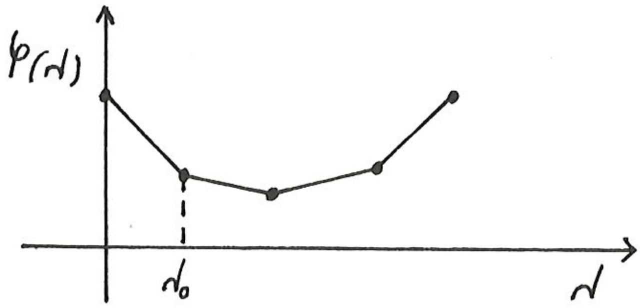


Figure 5

Choisissons alors pour λ_0, le premier point de non différentiabilité strictement positif. Dès lors, $f''(x_0; d) = 0$ et $f''(x_0; \cdot)$ constitue bien une fonction convexe. ∎

Proposition 2.2.4 : [Le cas différentiable]

Lorsque la fonction considérée est différentiable, $f''(x_0; \cdot)$ est convexe et continue.

Preuve :

Nous savons que

$$f''(x_0; d) = \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - f'(x_0; d) \right]$$

et, dans le cas envisagé, nous avons que

$$f''(x_0; d) = \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - \nabla f(x_0)^T d \right]$$

où $\nabla f(x_0)^T d$ est linéaire en d et

$$\frac{f(x_0 + \lambda .) - f(x_0)}{\lambda}$$

constitue une fonction convexe.

En effet, en utilisant la convexité de la fonction f, nous avons $\forall \mu \in [0, 1]$

$$\begin{aligned}
& \frac{f(x_0 + \lambda[\mu d_1 + (1-\mu) d_2]) - f(x_0)}{\lambda} \\
= \; & \frac{f[\mu(x_0 + \lambda d_1) + (1-\mu)(x_0 + \lambda d_2)] - \mu f(x_0) - (1-\mu) f(x_0)}{\lambda} \\
\leq \; & \frac{\mu f(x_0 + \lambda d_1) + (1-\mu) f(x_0 + \lambda d_2) - \mu f(x_0) - (1-\mu) f(x_0)}{\lambda} \\
= \; & \mu \frac{[f(x_0 + \lambda d_1) - f(x_0)]}{\lambda} + (1-\mu) \frac{[f(x_0 + \lambda d_2) - f(x_0)]}{\lambda}
\end{aligned}$$

Comme la limite de fonctions convexes reste convexe, nous obtenons la première partie du résultat.
La continuité de $f''(x_0; \cdot)$ s'établit facilement grâce au corollaire A.I.13. ∎

2.3 Développement de Taylor et méthode de Newton

Grâce à la dérivée directionnelle du second ordre, nous pouvons généraliser le développement de Taylor de la fonction f au cas non différentiable.
Nous obtenons ainsi par construction,

$$\forall\, \lambda > 0 \;\; f(x_0 + \lambda d) = f(x_0) + \lambda f'(x_0; d) + \frac{1}{2}\lambda^2 f''(x_0; d) + o(\lambda^2)$$

Dès lors, le problème d'optimisation (P) exposé dans le chapitre 0 se réécrit

$$(Q) \equiv \begin{cases} \min \quad f'(x_0; d) + \frac{1}{2} f''(x_0; d) \\ \text{sous } d \in I\!R^n \end{cases}$$

lorsque x_0 est pris comme point de départ.
La résolution du problème (Q) nous donne la direction d_0 et le point de l'"itération suivante est déterminé par

$$x_1 = x_0 + d_0$$

Cette façon de procéder présente cependant quelques inconvénients comme la possibilité de non-convergence de la méthode. Examinons ainsi l'exemple suivant.

Soit $f(x) = \max\{\frac{1}{2}(x-1)^2, \frac{1}{2}(x+1)^2\}$.

Prenons $x_0 > 0$ et représentons géométriquement la fonction ci-dessus

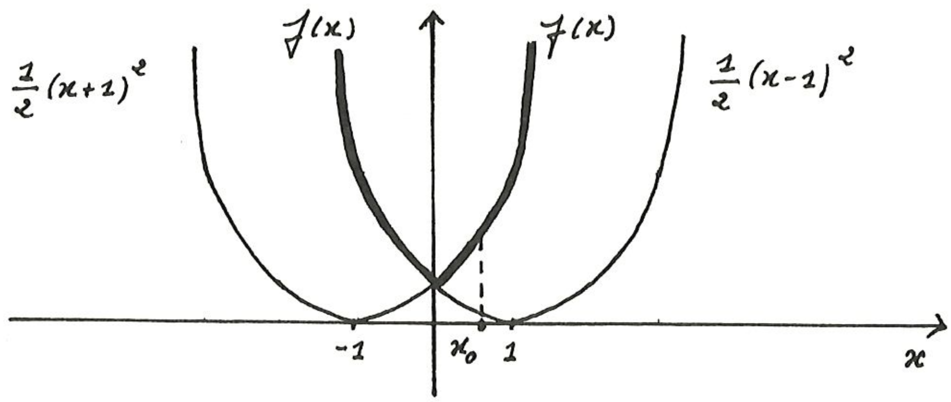


Figure 6

Pour obtenir le point x_1 de l'itération suivante, nous considérons l'approximation quadratique $\frac{1}{2}(x+1)^2$ de f. Celle-ci donne à x_1 la valeur -1. Nous répétons le processus et obtenons successivement $x_2 = 1, x_3 = -1, x_4 = 1, \cdots, x^k = (-1)^k, \cdots$, soit une suite de points oscillant entre -1 et 1. ∎

Un autre désavantage lié à cette méthode provient du fait que la fonction $f''(x_0; \cdot)$ n'est en général pas convexe, conne nous l'avons souligné précédemment. Dès lors, le problème (Q) devient difficile à résoudre ...

2.4 Dérivée directionnelle du second ordre et fonction d'appui

2.4.1 La notion de fonction d'appui

Soit un ensemble K convexe non vide inclus dans $\mathbb{R}^n$.
Nous rappelons que la fonction indicatrice de K, notée $\delta(K|\cdot)$ est une fonction convexe propre [cfr A.I.7] de $\mathbb{R}^n$ dans $\overline{\mathbb{R}}$ définie par

$$\delta(K|\cdot) = \begin{cases} 0 & \text{si } x \in K \\ +\infty & \text{sinon} \end{cases}$$

Elle constitue, en outre, une fonction semi-continue inférieurement lorsque K est fermé [cfr 4].
Considérons alors la conjuguée [cfr A.III.1] de l'indicatrice.
Nous obtenons ainsi la fonction $\delta^*(K|\cdot)$ t.q.

$$\forall\, y \in \mathbb{R}^n \quad \delta^*(K|y) = \sup_{x \in \mathbb{R}^n} \{< y, x > -\delta(K|x)\}$$

c'est-à-dire

$$\delta^*(K|y) = \sup_{x \in K} < y, x >$$

Cette fonction définie sur $\mathbb{R}^n$ et à valeurs dans $\overline{\mathbb{R}}$ porte le nom de *fonction d'appui* (ou *support*) de K. Nous pouvons en donner une interprétation géométrique.
Nous savons que les vecteurs x_1 et x_2 de $\mathbb{R}^n$, représentés ci-dessous, donnent le même produit scalaire avec y. Dès lors, en supposant que $||y|| = 1, \delta^*(K|y)$ correspond à la longueur du segment $0T$ ou encore, $\delta^*(K|y) =< y, x^* >$.

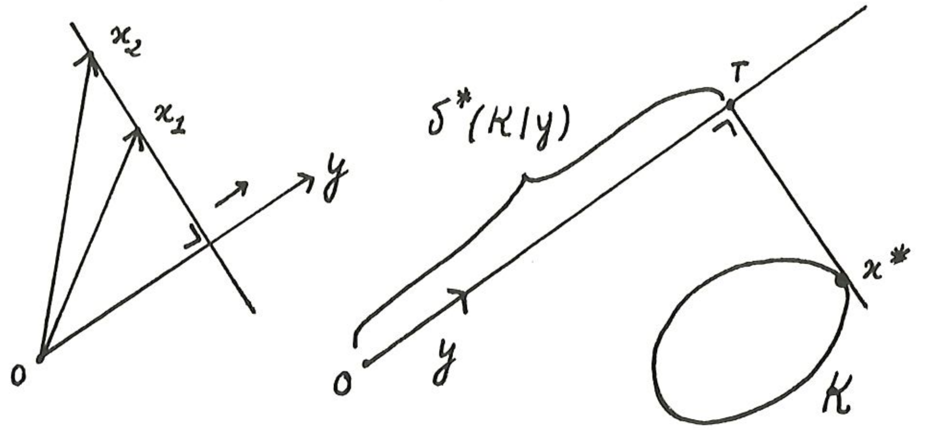


Figure 7

Nous avons immédiatement que,

K convexe fermé non vide
$\Rightarrow \delta^*(K|\cdot)$ convexe, propre, semi-continue inférieurement, positivement homogène de degré 1.

Nous énonçons à présent, un important théorème de Rockafellar (voir [8, proposition 13.2.1]).

Théorème 2.4.1.1 :

Soit g, une fonction convexe, propre semi-continue inférieurement, positivement homogène de degré 1.
Alors g est la fonction d'appui d'un convexe fermé, non vide K inclus dans $\mathbb{R}^n$.
De plus, $K = \{x^* \in \mathbb{R}^n | < x^*, d > \leq g(d) \quad \forall\, d \in \mathbb{R}^n\}$.

2.4.2 Propriétés de la forme quadratique $\mathrm{d} \to \mathrm{d}^{\mathrm{T}}\nabla^2\mathrm{f}(\mathrm{x_0})\mathrm{d}$ dans le cas deux fois différentiable

Nous savons que la matrice hessienne de f au point x_0, soit $\nabla^2 f(x_0)$, est symétrique et semi-définie positive. On peut dès lors lui associer un ellipsoïde [cfr A. VII]

$$E = [\nabla^2 f(x_0)]^{\frac{1}{2}}(\overline{B})$$

c'est-à-dire

$$E = \{\ [\nabla^2 f(x_0)]^{\frac{1}{2}}\ b \mid b \in \overline{B}\}$$

$\overline{B}$ désignant la boule unité fermée.
Dans le cas où $\nabla^2 f(x_0)$ est définie positive, donc inversible, on a aussi

$$E = \{y \in I\!R^n \mid y^T[\nabla^2 f(x_0)]^{-1} y \leq 1\}$$

En effet, nous prouvons le résultat plus général suivant.

Proposition 2.4.2.1 :

Si A est symétrique, inversible
Alors $A^{\frac{1}{2}}(\overline{B}) = \{y \mid y^T A^{-1} y \leq 1\}$

Démonstration :
Envisageons l'inclusion dans les deux sens.
Prenons $y = A^{\frac{1}{2}}z$ où $||z|| \leq 1$.
Nous avons alors, grâce au caractère symétrique de A,

$$\begin{aligned} y^T A^{-1} y &= z^T A^{\frac{1}{2}} A^{-1} A^{\frac{1}{2}} z \\ &= z^T z = ||z||^2 \leq 1 \end{aligned}$$

Considérons maintenant un élément y t.q. $y^T A^{-1} y \leq 1$
Nous cherchons donc à vérifier que

$$\exists z \text{ t.q. } ||z|| \leq 1 \text{ et } y = A^{\frac{1}{2}} z$$

Prenons alors $z = A^{-\frac{1}{2}}y$.
Nous obtenons ainsi

$$\begin{aligned} ||z||^2 &= ||A^{-\frac{1}{2}}y||^2 = y^T A^{-\frac{1}{2}} A^{-\frac{1}{2}} y \\ &= y^T A^{-1} y \leq 1 \end{aligned}$$

■

Nous établissons, à présent, le lien entre la forme quadratique $d \to d^T \nabla^2 f(x_0) d$ et la fonction d'appui de l'ellipsoïde E.

Proposition 2.4.2.2 :

$$\forall\, d \ \sqrt{d^T \nabla^2 f(x_0) d} = \sup_{e \in E} \ < e, d >$$

Démonstration :

Puisque $E = [\nabla^2 f(x_0)]^{\frac{1}{2}}(\overline{B})$, nous avons

$$\begin{aligned} \forall\, d \ \sup_{e \in E} \ < e, d > &= \sup_{||b|| \leq 1} < [\nabla^2 f(x_0)]^{\frac{1}{2}} b, d > \\ &= \sup_{||b|| \leq 1} \ d^T [\nabla^2 f(x_0)]^{\frac{1}{2}} b \end{aligned}$$

Comme le produit scalaire de deux vecteurs vaut le produit de leur norme par le cosinus de l'angle entre ces deux vecteurs, nous obtenons pour le vecteur b optimal,

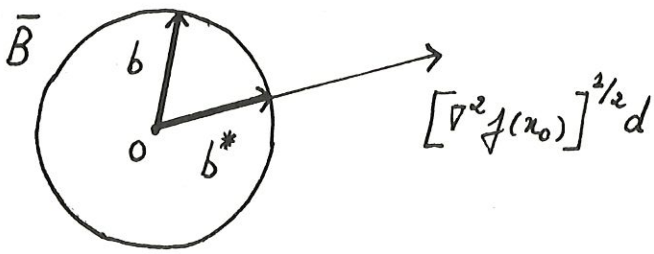


Figure 8

$$b^* = \frac{[\nabla^2 f(x_0)]^{\frac{1}{2}} d}{|| \ [\nabla^2 f(x_0)]^{\frac{1}{2}} d \ ||} = \frac{[\nabla^2 f(x_0)]^{\frac{1}{2}} d}{\sqrt{d^T \nabla^2 f(x_0) d}}$$

D'où,

$$\sup_{e\in E} < e, d > = \frac{d^T[\nabla^2 f(x_0)]^{\frac{1}{2}}[\nabla^2 f(x_0)]^{\frac{1}{2}} d}{\sqrt{d^T \nabla^2 f(x_0) d}} = \sqrt{d^T \nabla^2 f(x_0) d}$$

■

2.4.3 Dérivée directionnelle du second ordre et fonction d'appui

Nous souhaitons obtenir une propriété analogue à la proposition 2.4.2.2. dans le cas où $d^T\nabla^2 f(x_0)d$ est remplacé par $f''(x_0; d)$ (f n'étant pas nécessairement deux fois différentiable). Pour ce faire, nous allons tenter d'appliquer le théorème de Rockafellar dont l'énoncé est repris sous le nom de théorème 2.4.1.1.
Nous voulons exprimer $\sqrt{f''(x_0; d)}$ comme fonction d'appui d'un certain ensemble K convexe, fermé et non vide. Mais si $\sqrt{f''(x_0; d)}$ est bien positivement homogène et propre, elle n'est en général ni convexe, ni semi-continue inférieurement.
D'où les questions suivantes :

A) Comment convexifier $\sqrt{\mathbf{f''(x_0; d)}}$?

Par définition, nous avons

$$f''(x_0; d) = \lim_{\lambda\to 0+} \frac{2}{\lambda}\left[\frac{f(x_0+\lambda d) - f(x_0)}{\lambda} - f'(x_0; d)\right]$$

où $f'(x_0; \cdot)$ constitue une fonction convexe [voir A.IV.2] et $\dfrac{f(x_0+\lambda\cdot) - f(x_0)}{\lambda}$ est également convexe (puisque f l'est).
Comme la différence de deux fonctions convexes n'est pas nécessairement convexe, nous allons chercher à linéariser $f'(x_0; \cdot)$. Nous savons [cfr A.IV.5] que

$$f'(x_0; d) = \sup_{x\in\partial f(x_0)} < x, d >= \delta^*(\partial f(x_0)|d)$$

où $\partial f(x_0)$, sous-différentiel de f en x_0, est un ensemble convexe, fermé, borné, non vide [voir A.IV.1].
Soit x^* un point de $\partial f(x_0)$ satisfaisant $f'(x_0; d) =< x^*, d >$. Un tel point existe toujours puisque $\partial f(x_0)$ est compact.

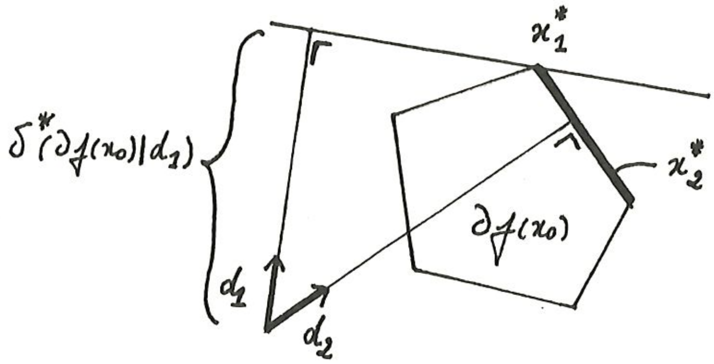


Figure 9

Inversément, nous nous posons la question de connaître $\{d \mid f'(x_0;d) =< x_0^*, d >\}$ pour x_0^* fixé dans $\partial f(x_0)$.

Dans le cas où $x_0^* \in$ intérieur de $\partial f(x_0)$, seule la direction nulle convient.

Si $x_0^* \in$ frontière de $\partial f(x_0)$, l'ensemble des directions répondant à la question est le cône normal [cfr A.V.1] $N(\partial f(x_0), x_0^*)$.

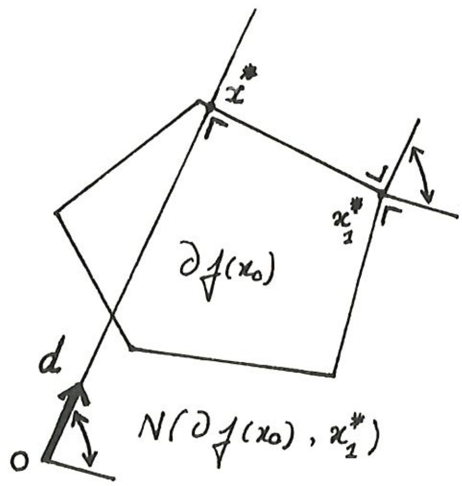


Figure 10

Ainsi, $f'(x_0;\cdot)$ est linéaire sur $N(\partial f(x_0), x_0^*)$ et l'on peut définir pour chaque $x_0^* \in \partial f(x_0)$

$$f''(x_0, x_0^*; d) = \begin{cases} \lim\limits_{\lambda \to 0^+} \frac{2}{\lambda}\left[\frac{f(x_0+\lambda d) - f(x_0)}{\lambda} - < x_0^*, d >\right] & \\ & \text{si } d \in N(\partial(x_0), x_0^*) \\ +\infty & \text{sinon} \end{cases}$$

En attribuant le nom de paire associée à tout couple (x_0^*, d) de $I\!R^n \times I\!R^n$ t.q. $d \in N(\partial f(x_0), x_0^*)$, on obtient encore

$$f''(x_0, x_0^*; d) = \begin{cases} f''(x_0; d) & \text{si } (x_0^*; d) \text{ est une paire associée} \\ +\infty & \text{sinon} \end{cases}$$

Nous présentons ci-dessous quelques propriétés de $f''(x_0, x_0^*; \cdot)$.

Proposition 2.4.3.1 :

a) *Si $x_0^* \in$ intérieur de $\partial f(x_0)$, alors $f''(x_0, x_0^*; \cdot) = \delta(0|\cdot)$*

b) *Lorsque f est différentiable au point $x_0, f''(x_0, \nabla f(x_0); \cdot) = f''(x_0; \cdot)$*

Démonstration :

a) Quand $x_0^* \in$ intérieur de $\partial f(x_0)$, seul le couple $(x_0^*, 0)$ constitue une paire associée. Comme $f''(x_0; 0) = 0$,

$$f''(x_0, x_0^*; d) = \begin{cases} 0 & \text{si } d = 0 \\ +\infty & \text{sinon} \end{cases}$$

b) Puisque f est différentiable au point x_0, le sous-différentiel $\partial f(x_0)$ se réduit au gradient $\nabla f(x_0)$ [cfr A.IV.4]. Dès lors, le cône normal en x_0^*

$$N(\partial f(x_0), x_0^*) = \{d \mid \ \forall\ x^* \in \partial f(x_0) = \{\nabla f(x_0)\} \quad 0 \geq < d, x^* - \nabla f(x_0) > \}$$

autrement dit, le couple (x_0^*, d) est une paire associée quelque soit la direction d envisagée. ∎

Proposition 2.4.3.2 :

a) $f''(x_0, x_0^*; \cdot) = f''(x_0; \cdot) + \delta(N(\partial f(x_0), x_0^*)|\cdot)$

b) $f''(x_0, x_0^*; 0) = 0$

c) $f''(x_0, x_0^*; d) \geq 0 \ \ \forall\, d$

Démonstration :

a) Il suffit d'envisager séparément les deux cas suivants :

$$d \in N(\partial f(x_0), x_0^*) \text{ et } d \notin N(\partial f(x_0), x_0^*)$$

b) preuve immédiate puisque $(x_0^*, 0)$ est une paire associée [voir proposition 2.2.1].

c) Ce résultat provient du fait que $f''(x_0; d)$ est déjà positive $\forall\, d$. ■

Proposition 2.4.3.3 :

$f''(x_0, x_0^*; .)$ *est convexe et positivement homogène de degré 2.*

Démonstration :

Attachons-nous tout d'abord à démontrer la convexité de $f''(x_0, x_0^*; .)$.
En vertu de la proposition A.I.9., il nous suffit de prouver que $f''(x_0, x_0^*; .)$ est propre, que son domaine [cfr A.I.6.] est convexe, et que

$$\forall\, d_1, d_2 \in \text{ dom } f''(x_0, x_0^*; .),\ \forall\, \mu \in [0,1] :$$

$$f''(x_0, x_0^*; \mu d_1 + (1-\mu) d_2) \leq \mu f''(x_0, x_0^*; d_1) + (1-\mu) f''(x_0, x_0^*; d_2)$$

Le caractère propre de $f''(x_0, x_0^*; .)$ s'établit immédiatement puisque cette fonction est positive et de valeur nulle en 0. Quant à la convexité du domaine, elle s'obtient tout aussi facilement en constatant que ce domaine correspond au cône normal $N(\partial f(x_0), x_0^*)$.
De plus, quand $d \in N(\partial f(x_0), x_0^*)$,

$$f'(x_0; d) = < x_0^*, d >$$

de sorte que

$$f''(x_0, x_0^*; .) = \lim_{\lambda \to 0+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda \cdot) - f(x_0)}{\lambda} - < x_0^*, . > \right]$$

Puisque $\dfrac{f(x_0 + \lambda \cdot) - f(x_0)}{\lambda}$ est convexe [cfr preuve de la proposition 2.2.4.], $f''(x_0, x_0^*; .)$ le devient aussi comme limite de fonctions convexes.
Par ailleurs, puisque $f''(x_0; .)$ est positivement homogène de degré 2, la thèse suit. ■

Rassemblant toutes les caractéristiques de $f''(x_0, x_0^*; d)$ nécessaires à l'application du théorème de Rockafellar, nous observons que $\sqrt{f''(x_0, x_0^*; d)}$ est bien convexe, propre, positivement homogène de degré un mais non nécessairement semi-continue inférieurement. En effet, reprenons l'exemple 1 avec $x_0 = (0,0)$ et $x_0^* = (0,0)$. Le dernier point est bien un élément de $\partial f(x_0)$ puisque

$$f(x) \geq f(0) + < 0, x > \quad \forall\, x \in I\!R^n .$$

Nous savons que le couple (x_0^*, d) constitue une paire associée si

$$d \in N(\partial f(x_0), x_0^*)$$

c'est-à-dire si

$$f'(x_0; d) = < x_0^*, d >$$

Or, pour $d = (d_1, d_2)$,

$$f'(x_0; d) = \begin{cases} 0 & \text{si } d_1 \leq 0 \\ 2d_1 & \text{si } d_1 > 0 \end{cases}$$

donc, $(x_0^*, d) = (0, d)$ sera paire associée pour $d_1 \leq 0$.
Dès lors,

$$\forall\, d = (d_1, d_2) \quad f''(x_0, x_0^*; d) = \begin{cases} 0 & \text{si } d_1 < 0 \\ 2d_2^2 & \text{si } d_1 = 0 \\ +\infty & \text{si } d_1 > 0 \end{cases}$$

puisque

$$f''(x_0, ; d) = \begin{cases} 0 & \text{si } d_1 < 0 \\ 2(d_1^2 + d_2^2) & \text{si } d_1 \geq 0 \end{cases}$$

Ainsi que nous l'affirme la proposition A.I.12., $f''(x_0, x_0^*; \cdot)$ est continue, donc semi-continue inférieurement, sur l'intérieur de $N(\partial f(x_0), x_0^*)$ mais, comme observé précédemment pour $f''(x_0; d)$, le caractére semi-continu inférieurement est mis en défaut sur la frontière du cône normal. ∎

B) Comment rendre $\sqrt{\mathbf{f''(x_0, x_0^*; d)}}$ semi-continue inférieurement ?

Nous rappelons que la fermeture d'une fonction convexe $g : I\!R^n \to I\!R \cup \{+\infty\}$ est la plus grande fonction semi-continue inférieurement qui minore g. On la note *cl* g.
D'un point de vue géométrique, *cl* g est la fonction dont l'épigraphe est $\overline{epi\ g}$.
Nous illustrons cette propriété au moyen de l'exemple suivant.

Soit une fonction $g : I\!R \to I\!R \cup \{+\infty\}$ t.q.

$$g(x) = \begin{cases} 0 & \text{si } x > 0 \\ +\infty & \text{si } x \leq 0 \end{cases}$$

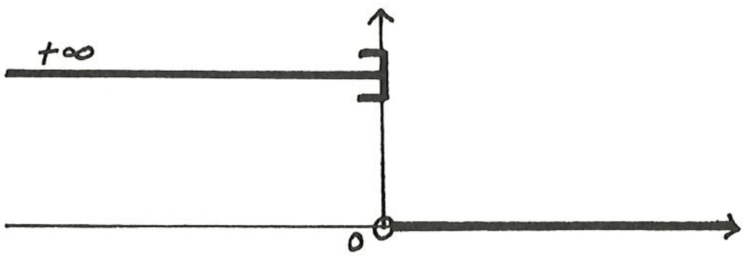


Figure 11

Nous avons

$$\begin{aligned} epi\ g &= \{(x,a) \in I\!R^n \times I\!R \mid g(x) \leq a\} \\ &= I\!R_0^+ \times I\!R_0 \\ \text{et} \quad \overline{epi\ g} &= I\!R^+ \times I\!R^+ \end{aligned}$$

Nous considérons à présent *cl* $g : I\!R \to I\!R \cup \{+\infty\}$ et obtenons

$$cl\ g(x) = \begin{cases} 0 & \text{si } x \geq 0 \\ +\infty & \text{si } x < 0 \end{cases}$$

avec $epi\ cl\ g(x) = I\!R^+ \times I\!R^+$. ∎

Proposition 2.4.3.4 :

La fermeture d'une fonction convexe, positive, positivement homogène est convexe, positive, positivement homogène et semi-continue inférieurement.

Démonstration :

Grâce aux propositions A.II.6 et A.II.4, il ne nous reste plus qu'à prouver les caractères positif et positivement homogène de $cl\ f$ où f désigne une fonction convexe, positive et positivement homogène de $I\!R^n$ dans $\overline{R}$.

a) $f \geq 0 \Rightarrow cl\ f \geq 0$:
En effet,

$$\begin{aligned} f \geq 0 &\Leftrightarrow epi(f) \subseteq I\!R^n \times I\!R^+ \\ &\Leftrightarrow epi(cl\ f) = \overline{epi(f)} \subseteq \overline{I\!R^n \times I\!R^+} = I\!R^n \times I\!R^+ \\ &\Leftrightarrow cl\ (f) \geq 0 \end{aligned}$$

b) f positivement homogène $\Rightarrow cl(f)$ positivement homogène :
De fait, par la proposition A.II.7, nous avons $\forall\ \alpha \geq 0$

$$\begin{aligned} cl\ f(\alpha x) &= \liminf_{y \to \alpha x} f(y) \\ &= \liminf_{z \to x} f(\alpha z) \\ &= \liminf_{z \to x} \alpha f(z) \end{aligned}$$

Pour clôturer la preuve, il nous suffit donc de montrer que nous pouvons permuter le scalaire α avec la limite inférieure.
Nous savons (cfr [8]) que

$$\liminf_{z \to x} \alpha f(z) = \lim_{\varepsilon \to 0^+} \varphi(\varepsilon) = \sup_{\varepsilon > 0} \varphi(\varepsilon)$$

où

$$\varphi(\varepsilon) = \inf\{\alpha f(z) \mid |z - x| \leq \varepsilon\}$$

Puisque le scalaire α peut "sortir" de l'infimum et du supremum, le résultat annoncé est maintenant prouvé. ∎

Dès lors, $cl\ \sqrt{f''(x_0, x_0^*; .)}(d)$ satisfait les hypothèses du théorème de Rockafellar et constitue donc la fonction d'appui d'un ensemble convexe, fermé, non vide, que nous noterons $\partial^2 f(x_0, x_0^*)$.

Ainsi, nous avons

$$cl\ \sqrt{f''(x_0, x_0^*; .)}(d) = \delta^*(\partial^2 f(x_0, x_0^*)|d) \tag{3}$$

Pour la commodité de l'écriture, nous admettrons dorénavant, l'abus de notation suivant

$$cl\ \sqrt{f''(x_0, x_0^*; .)}(d) = cl\ \sqrt{f''(x_0, x_0^*; d)}$$

CHAPITRE II

DIFFERENTIELS SECONDS

1 Différentiel second de f en $(\mathbf{x_0}, \mathbf{x_0^*})$

Nous savons déjà, en vertu du théorème I.2.4.1.1 et de l'égalité (3), que

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} \in I\!R^n \mid \; < x_0^{**}, d > \leq cl \sqrt{f''(x_0, x_0^*; d)} \quad \forall\, d \in I\!R^n\} \tag{4}$$

Cet ensemble portera le nom de *différentiel second* de la fonction f en (x_0, x_0^*).
Nous pouvons en donner une définition équivalente :

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} \in I\!R^n \mid \; < x_0^{**}, d > \leq \sqrt{f''(x_0, x_0^*; d)} \;\; \forall\, d \in I\!R^n\} \tag{5}$$

En effet, le sens (4) $\Rightarrow$ (5) est immédiat par application de la propriété A.II.6. D'autre part, par définition de la fermeture [cfr A.II.3.],

$$cl \sqrt{f''(x_0, x_0^*; d)} = \sup_{\substack{F \text{ affine} \\ F \leq \sqrt{f''(x_0, x_0^*; .)}}} F(d)$$

Or, $< x_0^{**}, . >$ est une fonction affine minorant $\sqrt{f''(x_0, x_0^*, .)}$ et donc,

$$< x_0^{**}, . > \leq \sup_{\substack{F \text{ affine} \\ F \leq \sqrt{f''(x_0, x_0^*; .)}}} F(.) \; = cl \sqrt{f''(x_0, x_0^*; d)}$$

■

Proposition 1.1 :

$\partial^2 f(x_0, x_0^*) = I\!R^n$ lorsque $x_0^* \in$ intérieur de $\partial f(x_0)$.

Démonstration :

Quand $x_0^* \in$ intérieur de $\partial f(x_0)$, le cône normal à $\partial f(x_0)$ au point x_0^* se réduit à la direction nulle. Dès lors,

$$f''(x_0, x_0^*; d) = \begin{cases} 0 & \text{si } d = 0 \\ +\infty & \text{sinon} \end{cases}$$

et le résultat annoncé est ainsi prouvé, étant donné (5). ∎

Proposition 1.2 :

$\partial^2 f(x_0, x_0^*)$ est un ensemble convexe, fermé, contenant le cône tangent $T(\partial f(x_0), x_0^*)$ à $\partial f(x_0)$ au point x_0^*.

Démonstration :

Notons que les caractères convexe et fermé de $\partial^2 f(x_0, x_0^*)$ sont implicitement réalisés par définition ($\partial^2 f(x_0, x_0^*) = K$ dans le théorème I.2.4.1.1). Nous n'en faisons ici que la vérification.

Prouver la convexité du différentiel du second ordre revient à considérer deux points x_0^{**} et $y_0^{**} \in \partial^2 f(x_0, x_0^*)$ et à vérifier que

$$\forall\ \lambda \in [0,1]\ \ \lambda x_0^{**} + (1-\lambda) y_0^{**} \in \partial^2 f(x_0, x_0^*)$$

c'est-à-dire

$$\forall\ \lambda \in [0,1]\ \ < \lambda x_0^{**} + (1-\lambda) y_0^*, d > \leq \sqrt{f''(x_0, x_0^*; d)}\ \ \forall\ d$$

Ceci s'obtient par la propriété de linéarité du produit scalaire.
Etablissons à présent, le caractère fermé de $\partial^2 f(x_0, x_0^*)$. Pour ce faire, prenons une suite de points $(x_n^{**})_n$ dans $\partial^2 f(x_0, x_0^*)$, convergente vers un élément x^{**} et montrons que ce dernier appartient à $\partial^2 f(x_0, x_0^*)$. Autrement dit, vérifions que

$$< x^{**}, d > \leq \sqrt{f''(x_0, x_0^*; d)}\ \ \forall\ d$$

Or,

$$< x_n^{**}, d > \leq \sqrt{f''(x_0, x_0^*; d)}\ \ \forall\ d$$

et

$$< x_n^{**}, d > = < x_n^{**} - x^{**}, d > + < x^{**}, d >$$

Dès lors, par continuité du produit scalaire,

$$\begin{aligned} < \lim_{n\to\infty} (x_n^{**} - x^{**}), d > + < x^{**}, d > &= 0 + < x^{**}, d > \\ &\leq \sqrt{f''(x_0, x_0^*; d)} \ \forall\, d \end{aligned}$$

Montrons finalement que

$$T(\partial f(x_0), x_0^*) \subseteq \partial^2 f(x_0, x_0^*)$$

Nous avons

$$cl\ \sqrt{f''(x_0, x_0^*; .)} \geq \delta(N(\partial f(x_0), x_0^*) \mid .) \tag{6}$$

et

$$\begin{aligned} \delta(N(\partial f(x_0), x_0^*) \mid .) &= \delta^*(T(\partial f(x_0), x_0^*) \mid .) \\ &= \sup_{x \in T(\partial f(x_0), x_0^*)} < ., x > \end{aligned} \tag{7}$$

En effet, (6) s'obtient en séparant les cas :

(i) Si $d \in N(\partial f(x_0), x_0^*)$ c'est-à-dire (x_0^*, d) constitue une paire associée. Dans cette éventualité, le membre de droite de l'inégalité s'annule et le membre de gauche est positif.

(ii) Si $d \notin N(\partial f(x_0), x_0^*)$, chacun des deux membres de l'inégalité prend une valeur infinie.

Nous établissons l'égalité (7) en procédant de la même façon :

(i) Si $d \in N(\partial f(x_0), x_0^*)$, alors

$$\forall\, x \in T(\partial f(x_0), x_0^*) \quad < d, x > \leq 0$$

puisque $T(\partial f(x_0), x_0^*)^- = N(\partial f(x_0), x_0^*) \ \forall\, x_0^* \in \partial f(x_0)$ [cfr A.V.2]
Or, $0 \in N(\partial f(x_0), x_0^*)$, donc

$$\sup_{x \in T(\partial f(x_0), x_0^*)} < d, x > = 0$$

(ii) Si $d \notin N(\partial f(x_0), x_0^*)$ c'est-à-dire si $d \notin T(\partial f(x_0), x_0^*)^-$, alors $\exists x^* \in T(\partial f(x_0), x_0^*)$ tel que $< d, x^* > > 0$.
Mais si x^* appartient au cône tangent, λx^* y appartient également pour tout $\lambda > 0$.
Dès lors,

$$\sup_{\lambda > 0} < d, \lambda x^* > \ \leq \sup_{x \in T(\partial f(x_0), x_0^*)} < d, x >$$

ou encore

$$\sup_{\lambda > 0} \lambda < d, x^* > \ \leq \sup_{x \in T(\partial f(x_0), x_0^*)} < d, x >$$

Nous concluons au moyen de la proposition A.VIII.1 appliquée aux deux convexes fermés suivants :

$$\partial^2 f(x_0, x_0^*) \text{ et } T(\partial f(x_0), x_0^*)$$

■

Penchons-nous à nouveau sur l'exemple 1 afin de constater que $\partial^2 f(x_0, x_0^*)$ n'est pas nécessairement borné.
Puisque, pour $x_0 = (0,0)$ et $x_0^* = (0,0)$,

$$f''(x_0, x_0^*; d) = \begin{cases} 0 & \text{si } d_1 < 0 \\ 2d_2^2 & \text{si } d_1 = 0 \\ +\infty & \text{si } d_1 > 0 \end{cases}$$

nous obtenons

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} \mid < x_0^{**}, d > \le \begin{cases} 0 & \text{si } d_1 < 0 \\ \sqrt{2d_2^2} & \text{si } d_1 = 0 \\ +\infty & \text{si } d_1 > 0 \end{cases} \}$$

Les deux premières conditions peuvent se réécrire

$$\begin{aligned} x_{01}^{**}d_1 + x_{02}^{**}d_2 &\le 0 \quad \text{si } d_1 < 0 \; \forall \; d_2 & (8) \\ x_{02}^{**}d_2 &\le \sqrt{2d_2^2} \; \forall \; d_2 & (9) \end{aligned}$$

avec $x_0^{**} = (x_{01}^{**}, x_{02}^{**})$.
Nous observons alors que (8) implique

$$x_{01}^{**} \ge 0 \text{ et } x_{02}^{*} = 0$$

(considérer respectivement $d_2 = 0$ puis, d_2 de très grande valeur).
Ainsi, (9) est toujours vérifiée.
Finalement, nous pouvons identifier $\partial^2 f(x_0, x_0^*)$ à

$$\{x_0^{**} \mid x_{01}^{**} \ge 0 \text{ et } x_{02}^{**} = 0\}$$

c'est-à-dire $\partial^2 f(x_0, x_0^*) = I\!R^n \times \{0\}$ qui n'est évidemment pas borné. ■

2 Différentiel second de f en $\mathbf{x_0}$

Nous définirons le *différentiel second* de f en x_0 par

$$\partial^2 f(x_0) = \bigcap_{x_0^* \in \partial f(x_0)} \partial^2 f(x_0, x_0^*) \tag{10}$$

En fait, puisque $\partial^2 f(x_0, x_0^*)$ vaut l'espace tout entier lorsque x_0^* se trouve dans l'intérieur de $\partial f(x_0)$, seuls les x_0^* de la frontière de $\partial f(x_0)$ joueront un rôle significatif dans la définition donnée ci-dessus.

Nous pourrions également, restreindre l'ensemble sur lequel porte l'intersection en ne considérant que les points extrémaux de $\partial f(x_0)$.

En effet, il suffit de démontrer l'inclusion non triviale

$$\bigcap_{x_0^* \in \text{ Extr } \partial f(x_0)} \partial^2 f(x_0, x_0^*) \subseteq \partial^2 f(x_0)$$

Puisque $\partial f(x_0)$ est l'enveloppe convexe de ses points extrémaux, tout élément x_0^* de cet ensemble pourra s'exprimer comme combinaison convexe des points extrémaux $x_0^{*1}, \cdots, x_0^{*k}$. Ainsi,

$$x_0^* = \sum_{i=1}^{k} \alpha_i x_0^{*i}$$

avec

$$\sum_{i=1}^{k} \alpha_i = 1 \text{ et } \alpha_i > 0 \;\forall\, i$$

Considérons maintenant $d \in N(\partial f(x_0), x_0^*)$ de sorte que

$$< x_0^*, d >= \sum_{i=1}^{k} \alpha_i < x_0^{*i}, d >= f'(x_0; d)$$

Comme $x_0^{*i} \in \partial f(x_0) \;\forall\, i$, nous obtenons [cfr A.IV.3]

$$< x_0^{*i}, d >= f'(x_0; d) \;\;\forall\, i$$

ce qui signifie que $\alpha \in N(\partial f(x_0), x_0^{*i}) \;\forall\, i$.

Par conséquent, nous avons prouvé que

$$N(\partial f(x_0), x_0^*) \subseteq N(\partial f(x_0), x_0^{*i}) \;\;\forall\, i$$

Dès lors,

$$f''(x_0, x_0^*; .) \geq f''(x_0, x_0^{*i}; .) \;\;\forall\, i$$

ce qui implique par (5) que

$$\partial^2 f(x_0, x_0^*) \supseteq \partial^2 f(x_0, x_0^{*i}) \quad \forall\, i$$

L'inclusion annoncée est ainsi démontrée. ∎

Nous pouvons donner du différentiel second de f en x_0 une définition équivalente.

Proposition 2.1 :

$$\partial^2 f(x_0) = \{x^* \in I\!R^n \mid \; < x^*, d > \leq \sqrt{f''(x_0; d)} \;\; \forall\, d\}$$

Démonstration :

Par les définitions (10) et (5), nous savons que $x_0^{**} \in \partial^2 f(x_0)$ si et seulement si

$$< x_0^{**}, d > \leq \inf_{x_0^* \in \partial f(x_0)} \sqrt{f''(x_0, x_0^*; d)} \;\; \forall\, d$$

D'autre part,

$$\inf_{x_0^* \in \partial f(x_0)} \sqrt{f''(x_0, x_0^*; d)} = \sqrt{f''(x_0; d)}$$

de telle sorte que

$$x_0^{**} \in \partial^2 f(x_0) \quad \Leftrightarrow \; < x_0^{**}, d > \leq \sqrt{f''(x_0; d)} \;\; \forall\, d$$

∎

Nous prouvons à présent quelques propriétés fondamentales du différentiel second de f en x_0.

Proposition 2.2 :

$\partial^2 f(x_0)$ est un ensemble convexe, compact, contenant l'origine.

Démonstration :

Considérons deux éléments de $\partial^2 f(x_0)$, soient x_0^{**} et y_0^{**}.
Nous désirons prouver que $\lambda x_0^{**} + (1-\lambda) y_0^{**} \in \partial^2 f(x_0)$, c'est-à-dire que

$$\forall\, x_0^* \in \partial f(x_0) \quad \lambda x_0^{**} + (1-\lambda) y_0^{**} \in \partial^2 f(x_0, x_0^*)$$

Or, $\forall\ x_0^* \in \partial f(x_0)$

$$x_0^{**} \text{ et } y_0^{**} \in \partial^2 f(x_0, x_0^*)$$

D'où, par convexité de ce dernier ensemble [cfr proposition 1.2], la première partie du résultat est obtenue.

Montrons à présent le caractère compact de $\partial^2 f(x_0)$.

Il s'agit tout d'abord d'un ensemble fermé puisque défini par une intersection de fermés.

Prouvons maintenant que $\partial^2 f(x_0)$ est borné. Prenons pour cela x_0^{**} dans $\partial^2 f(x_0)$, nous avons alors

$$\begin{aligned} ||x_0^{**}|| &= \sup_{||d||=1} \ < x_0^{**}, d > \\ &\leq \sup_{||d||=1} \sqrt{f''(x_0; d)} \end{aligned}$$

par la proposition 2.1.

D'autre part, au moyen de (2), nous pouvons écrire

$\forall\ \varepsilon > 0\ \ \exists \lambda_0 > 0$ tel que

$$\left| f''(x_0; d) - \frac{1}{\lambda_0}[f'(x_0 + \lambda_0 d; d) - f'(x_0; d)] \right| < \varepsilon$$

c'est-à-dire

$$|f''(x_0; d)| < \varepsilon + \frac{1}{\lambda_0}|f'(x_0 + \lambda_0 d; d)| + \frac{1}{\lambda_0}|f'(x_0; d)]$$

Or, par la proposition A.IV.5 et le caractère borné de $\partial f(x_0)$, nous avons

$$\begin{aligned} f'(x_0; d) &= \max_{x_0^* \in \partial f(x_0)} \ < x_0^*, d > \\ &\leq K\ ||d|| = K \end{aligned}$$

De même, nous obtenons

$$\begin{aligned} f'(x_0 + \lambda_0 d; d) &= \max_{x_0^* \in \partial f(x_0 + \lambda_0 d)} \ < x_0^*, d > \\ &\leq K_1\ \ ||d|| = K_1 \end{aligned}$$

Consécutivement,

$$|f''(x_0; d)| < \varepsilon + \frac{1}{\lambda_0}(|K| + |K_1|)$$

et ainsi $\partial^2 f(x_0)$ devient borné.

Prouver que l'origine est un élément de $\partial^2 f(x_0)$ ne pose aucune difficulté puisque $\forall\ x_0^* \in \partial f(x_0) \quad 0 \in \partial^2 f(x_0, x_0^*)$. ∎

Proposition 2.3 :

La fonction d'appui du différentiel second de f en x_0 est la biconjuguée de $\sqrt{f''(x_0;\cdot)}$.

Démonstration :

Montrons tout d'abord le résultat préliminaire suivant.
Si $g : I\!R^n \to I\!R \cup \{+\infty\}$ est une fonction positivement homogène, alors $g^* = 0$ sur dom (g^*).
En effet, pour tout $\lambda > 0, \lambda \neq 1$ et tout $y \in I\!R^n$, nous avons

$$g^*(y) = \sup_{x \in I\!R^n} \{< x, y > -g(x)\}$$

et donc,

$$\begin{aligned} \lambda g^*(y) &= \sup_{x \in I\!R^n} \{\lambda < x, y > -\lambda g(x)\} \\ &= \sup_{x \in I\!R^n} \{< \lambda x, y > -g(\lambda x)\} \\ &= \sup_{z \in I\!R^n} \{< z, y > -g(z)\} \\ &= g^*(y) \end{aligned}$$

D'où,

$$g^*(y) = \begin{cases} 0 & \text{si} \quad y \in \text{dom } (g^*) \\ +\infty & \text{sinon} \end{cases}$$

Nous pouvons déduire de cette propriété le caractère positivement homogène de la biconjuguée de $\sqrt{f''(x_0; d)}$.
De fait, notant

$$g(.) = \sqrt{f''(x_0; .)}$$

(il s'agit bien d'une fonction positivement homogène [voir proposition I.2.2.2]), nous obtenons

$$g^{**}(\lambda x) = \sup_{y \in I\!R^n} \{< y, \lambda x > -g^*(y)\} = \sup_{y \in \text{dom } (g^*)} < y, \lambda x >$$

et

$$\lambda g^{**}(x) = \sup_{y \in I\!R^n} \{< y, \lambda x > -\lambda g^*(y)\} = \sup_{y \in \text{dom } (g^*)} < y, \lambda x >$$

Par ailleurs, nous savons, grâce à la proposition A.III.2, que $(\sqrt{f''(x_0;.)})^{**}$ constitue une fonction convexe, propre, semi-continue inférieurement.
Dès lors, nous pouvons appliquer le théorème I.2.4.1.1 afin de déduire que

$$(\sqrt{f''(x_0;.)})^{**} = \delta^*(K|.)$$

où

$$K = \{x_0^{**} \in I\!R^n \mid < x_0^{**}, d > \leq (\sqrt{f''(x_0;.)})^{**} \ \forall\, d \in I\!R^n\}$$

Ce dernier ensemble n'est autre que $\partial^2 f(x_0)$.
En effet, nous pouvons facilement prouver que

$$\{x_0^{**} \in I\!R^n \mid < x_0^*, d > \leq \sqrt{f''(x_0;d)} \ \forall\, d \in I\!R^n\} \supseteq K$$

en faisant appel à la propriété suivante :

$$f^{**} \leq f \ \forall\, f \qquad \text{[cfr 4]}$$

L'autre inclusion s'établit ensuite, en considérant

$$\begin{aligned} < x_0^{**}, d > \quad & \leq \sqrt{f''(x_0;.)} \\ \Rightarrow \quad (< x_0^{**}, . >)^{**} \quad & \leq (\sqrt{f''(x_0;.)})^{**} \qquad \text{[cfr 4]} \end{aligned}$$

Nous pouvons alors conclure au moyen de la proposition A.III.4 puisque $< x_0^{**}, . >$ est convexe, propre, semi-continue inférieurement. ∎

Proposition 2.4 :

$\partial^2 f(x_0) + T(\partial f(x_0), x_0^*) \subseteq \partial^2 f(x_0, x_0^*)$.
L'égalité s'obtient lorsque $f''(x_0;.)$ est convexe.

Démonstration :

Soit $z^{**} \in \partial^2 f(x_0) + T(\partial f(x_0), x_0^*)$. Autrement dit, $z^{**} = z_1^{**} + z_2^{**}$ avec

$$z_1^{**} \in \partial^2 f(x_0) \text{ et } z_2^{**} \in T(\partial f(x_0), x_0^*) = N(\partial f(x_0), x_0^*)$$

Ainsi,

$$< z_1^{**}, d > \leq \sqrt{f''(x_0;d)} \ \forall\, d \in I\!R^n$$

et

$$< z_2^{**}, d > \leq \delta(N(\partial f(x_0), x_0^*)|d) \ \ \forall\ d \in I\!R^n$$

Additionnons maintenant membre à membre ces deux inégalités. Nous obtenons

$$\begin{aligned} < z^{**}, d > &\leq \sqrt{f''(x_0; d)} + \delta(N(\partial f(x_0), x_0^*)|d) \\ &= \sqrt{f''(x_0, x_0^*; d)} \ \ \forall\ d \in I\!R^n \end{aligned}$$

c'est-à-dire $z^{**} \in \partial^2 f(x_0, x_0^*)$ [cfr (5)].
Par ailleurs,

$$\begin{aligned} & x_0^{**} \in \partial(\sqrt{f''(x_0; .)} + \delta(N(\partial, f(x_0), x_0^*)|.)(0) \\ \Leftrightarrow \quad & \sqrt{f''(x_0; x)} + \delta(N(\partial f(x_0), x_0^*)|x) \geq 0+ < x_0^{**}, x > \ \ \forall\ x \in I\!R^n \end{aligned}$$

expression encore équivalente [cfr proposition I.2.4.3.2.a.] à

$$< x_0^{**}, x > \leq \sqrt{f''(x_0, x_0^*; x)} \ \ \forall\ x \in I\!R^n$$

c'est-à-dire $x_0^{**} \in \partial^2 f(x_0, x_0^*)$.
Ainsi,

$$\partial^2 f(x_0, x_0^*) = \partial(\sqrt{f''(x_0; .)} + \delta(N(\partial f(x_0), x_0^*)|.)(0)$$

Ajoutons à présent l'hypothèse de convexité de $f''(x_0; .)$ pour obtenir

$$\partial^2 f(x_0, x_0^*) \subseteq \partial\sqrt{f''(x_0; .)}(0) + \partial\delta(N(\partial f(x_0), x_0^*)|.)(0) \tag{11}$$

au moyen de la proposition A.IV.6.a. et A.I.13.
De plus, nous établissons que
$\partial\sqrt{f''(x_0; .)}(0) + \partial\delta(N(\partial f(x_0), x_0^*)(.)(0)$

$$\begin{aligned} &= \{x_0^* \in I\!R^n \mid \sqrt{f''(x_0, x)} \geq 0+ < x_0^*, x > \ \ \forall\ x \in I\!R^n\} \\ &\quad +N(N(\partial f(x_0), x_0^*), 0) \end{aligned} \tag{12}$$

par définition du sous-différentiel et du cône normal.
Finalement, (11) et (12) permettent l'obtention de l'inclusion souhaitée, à savoir,

$$\partial^2 f(x_0, x_0^*) \subseteq \partial^2 f(x_0) + T(\partial f(x_0), x_0^*)$$

■

Proposition 2.5 :

Quand f est deux fois différentiable,

$$\partial^2 f(x_0) = E = [\nabla^2 f(x_0)]^{1/2}(\bar{B})$$

Démonstration :

Puisque $\partial^2 f(x_0)$ et l'ellipsoïde E sont des parties convexes fermées, nous obtiendrons la thèse en prouvant l'égalité des fonctions d'appui [cfr A.VIII.1].
Nous savons déjà, en vertu de sa proposition 2.3, que

$$\delta^*(\partial^2 f(x_0)|d) = (\sqrt{f''(x_0;d)})^{**} \ \forall\ d \in I\!R^n$$

Nous réécrivons alors cette égalité en tenant compte du caractère deux fois différentiable de f

$$\delta^*(\partial^2 f(x_0)|d) = (\sqrt{d^T\nabla^2 f(x_0)d})^{**}$$

Puisque $d \to \sqrt{d^T\nabla^2 f(x_0)d}$ constitue une fonction convexe, propre et semi-continue inférieurement [voir proposition I.2.2.4], nous pouvons déduire [cfr A.III.4] que

$$\delta^*(\partial^2 f(x_0)|d) = \sqrt{d^T\nabla^2 f(x_0)d)} \ \forall\ d \in I\!R^n$$

D'autre part, nous avons déjà établi [cfr proposition 2.4.2.2] que

$$\sup_{e\in E} < e, d >= \sqrt{d^T\nabla^2 f(x_0)d)} \ \forall\ d \in I\!R^n$$

Ainsi,

$$\delta^*(\partial^2 f(x_0)|d) = \delta^*(E|d)$$

■

Nous examinons maintenant quelques exemples .
Le premier reprend la fonction déjà considérée dans l'exemple 1.
Soit donc $f(x_1, x_2) = \max(0, x_0^2 + x_2^2 + 2x_1)$.
Pour $x_0 = (0,0)$, nous obtenons, au moyen de la proposition A.IV.6.b., que

$$\partial f(x_0) = Co\{\bigcup_{i\in I(x_0)} \partial f_i(x_0)\}$$

où

$$I(x_0) = \{i = 1,2 \mid f_i(x_0) = \sup_{1 \leq j \leq 2} f_j(x_0)\}$$

c'est-à-dire

$$\partial f(x_0) = Co\{\bigcup_{i=1,2} \partial f_i(x_0)\}$$

avec

$$\partial f_1(x_0) = \{\nabla f_1(0,0)\} = (0,0)$$

et

$$\partial f_2(x_0) = \{\nabla f_2(0,0)\} = (2,0)$$

puisque f_1 et f_2 sont des fonctions différentiables.
Dès lors, nous envisageons les trois hypothèses suivantes pour le sous-gradient.

a) $x_0^* = (0,0)$:

Ce cas a déjà été étudié et nous recopions le résultat obtenu.

$$\partial^2 f(x_0, x_0^*) = I\!R^+ \times \{0\}$$

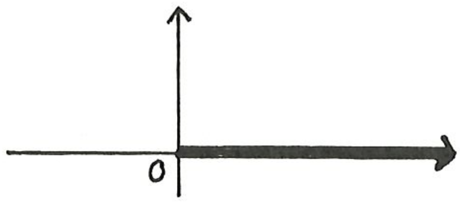


Figure 12

Nous remarquons que dans ce cas précis, nous obtenons l'égalité entre $\partial^2 f(x_0, x_0^*)$ et le cône tangent $T(\partial f(x_0), x_0^*)$ [voir proposition 1.2].

b) $x_0^* = (2,0)$:

Nous avons établi précédemment que

$$f''(x_0; d) = \begin{cases} 0 & \text{si } d_1 < 0 \\ 2(d_1^2 + d_2^2) & \text{si } d_1 \geq 0 \end{cases}$$

Par ailleurs, nous avons que

$$(x_0^*, d) \text{ constitue une paire associée } \Leftrightarrow \; < x_0^*, d >= f'(x_0; d)$$

ou encore, grâce à la proposition A.IV.7,

$$< x_0^*, d >= \max(\nabla f_1(x_0)^T d, \nabla f_2(x_0)^T d)$$

c'est-à-dire

$$2d_1 + 0 = \max(0, 2d_1)$$

Ainsi, (x_0^*, d) sera paire associée pour $d_1 \geq 0$.
Consécutivement,

$$f''(x_0, x_0^*; d) = \begin{cases} 2(d_1^2 + d_2^2) & \text{si } d_1 \geq 0 \\ +\infty & \text{si } d_1 < 0 \end{cases}$$

et

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} | < x_0^{**}, d > \leq \left\{ \begin{array}{ll} \sqrt{2(d_1^2 + d_2^2)} & \text{si } d_1 > 0 \\ \sqrt{2d_2^2} & \text{si } d_1 = 0 \\ +\infty & \text{si } d_1 < 0 \end{array} \right\} \forall\, d_2\}$$

Nous réécrivons ces conditions pour $x_0^{**} = (x_{01}, x_{02})$

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} | \left\{ \begin{array}{rcll} x_{01}d_1 + x_{02}d_2 & \leq & \sqrt{2d_1^2 + 2d_2^2} & \text{si } d_1 > 0 \; \forall\, d_2 \\ x_{02}d_2 & \leq & \sqrt{2}|d_2| & \forall\, d_2 \end{array} \right\}\}$$

Prouvons que ce dernier ensemble n'est autre que l'union disjointe

$$A_1 \dot{\cup} A_2$$

avec

$$A_1 = \{x_0^{**} = (x_{01}, x_{02}) \mid x_{01} \leq 0 \text{ et } x_{02}d_2 \leq \sqrt{2}d_2\}$$

et

$$A_2 = \{x_0^{**} = (x_{01}, x_{02}) \mid x_{01} > 0 \text{ et } x_{01}^2 + x0_2{}^2 - 2 \leq 0\}$$

Considérons premièrement l'inclusion

$$\partial^2 f(x_0, x_0^*) \subseteq A_1 \dot{\cup} A_2$$

Prenons x_0^{**} dans $\partial^2 f(x_0, x_0^*)$ et envisageons les deux possibilités suivantes

1^0) $x_{01} \leq 0$:

Dans ce cas, x_0^{**} appartient à A_1 et donc à $A_1 \dot{\cup} A_2$.

2^0) $x_{01} > 0$:

Posons alors $d_1 = x_{01} > 0$ et $d_2 = x_{02}$. Nous avons

$$x_{01}^2 + x_{02}^2 \leq \sqrt{2} \quad \sqrt{x_{01}^2 + x_{02}^2}$$

c'est-à-dire

$$0 \leq \sqrt{x_{01}^2 + x_{02}^2} \leq \sqrt{2}$$

Ainsi, $x_0^{**} \in A_2$ et donc à $A_1 \dot{\cup} A_2$.

Démontrons à présent, l'autre inclusion.

Pour ce faire, considérons un élément x_0^{**} de $A_1 \dot{\cup} A_2$. Deux cas peuvent alors se présenter.

1^0) $x_{01} \leq 0$:

Sous cette condition, nous avons que $x_0^{**} \in A_1$. Dès lors, nous trouvons déjà que $x_{02} d_2 \leq \sqrt{2} d_2$.

D'autre part,

$$\begin{aligned} x_{01} d_1 + x_{02} d_2 \leq x_{02} d_2 &\leq \sqrt{2} \ d_2 \\ &\leq \sqrt{2} \ |d_2| = \sqrt{2} \ \sqrt{d_2^2} \\ &\leq \sqrt{2} \ \sqrt{d_1^2 + d_2^2} \end{aligned}$$

Cette dernière inégalité étant vraie $\forall \ d_1$, le reste $\forall \ d_1 > 0$.

2^0) $x_{01} > 0$:

Dans ce cas, nous pouvons affirmer que $x_0^{**} \in A_2$ c'est-à-dire que $x_0^{**} \in \sqrt{2}\bar{B}$.

Considérons tout d'abord

$$< x_0^{**}, d > = x_{01} d_1 + x_{02} d_2 \leq \ ||x_0^{**}|| \ \ ||d||$$

par l'inégalité de Cauchy-Schwarz.

Dès lors, puisque $x_0^{**} \in \sqrt{2}\bar{B}$, nous avons que

$$\begin{aligned} x_{01} d_1 + x_{02} d_2 &\leq \sqrt{2} \ ||d|| \\ &= \sqrt{2} \ \sqrt{d_1^2 + d_2^2} \end{aligned}$$

Cette inégalité est vérifiée $\forall\ d_1$ et donc $\forall\ d_1 > 0$.
Par ailleurs, nous pouvons affirmer que

$$-\sqrt{2} \leq x_{02} \leq \sqrt{2}$$

en utilisant à nouveau l'hypothèse $x_0^{**} \in \sqrt{2}\bar{B}$.
De là, nous déduisons,

(i) Si $d_2 \leq 0$:

$$x_{02}d_2 \leq -\sqrt{2}\ \ d_2 = \sqrt{2}\ \ |d_2|$$

(ii) Si $d_2 > 0$:

$$x_{02}d_2 \leq \sqrt{2}\ \ d_2 - \sqrt{2}\ |\ d_2|$$

L'égalité annoncée se trouve ainsi vérifiée.
Dès lors, nous obtenons la représentation géométrique suivante.

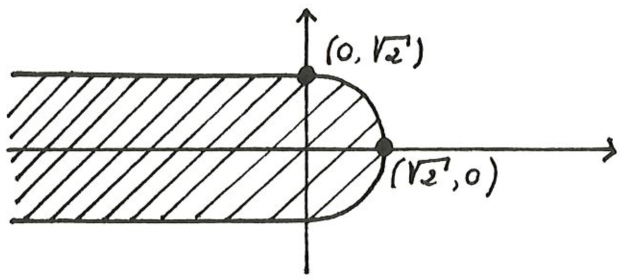


Figure 13

c) $x_0^* = (\alpha, 0)$ avec $\alpha \in]0, 2[$:
Sous cette hypothèse, nous obtenons que

$$\begin{aligned}
(x_0^*, d) \text{ est une paire associée} \quad &\Leftrightarrow \quad < x_0^*, d >= f'(x_0; d) = \max(0, 2d_1) \\
&\Leftrightarrow \quad \alpha d_1 = \max(0, 2d_1)\ \ \forall\ \alpha \in]0, 2[\\
&\Leftrightarrow \quad d_1 = 0
\end{aligned}$$

Dès lors,

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} | < x_0^{**}, d > \leq \left\{ \begin{array}{ll} \sqrt{2d_2^2} & \text{si } d_1 = 0 \\ +\infty & \text{si } d_1 \neq 0 \end{array} \right\} \ \forall\ d_2\}$$

Nous explicitons la première contrainte pour $x_0^{**} = (x_{01}, x_{02})$

$$x_{02}d_2 \leq \sqrt{2}|d_2|$$

Nous pouvons ainsi conclure que l'ensemble

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} = (x_{01}, x_{02}) \mid -\sqrt{2} \leq x_{02} \leq \sqrt{2}\}$$

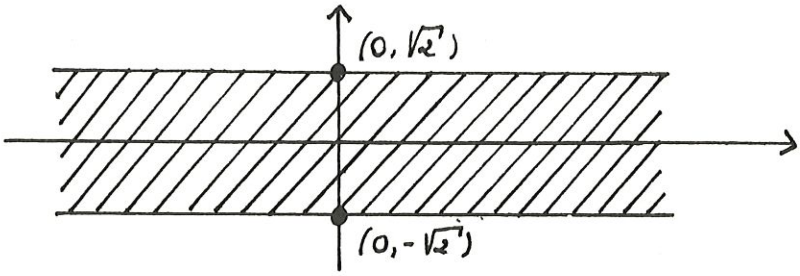


Figure 14

Consécutivement, nous trouvons au moyen de (10) que

$$\partial^2 f(x_0) = [0, \sqrt{2}] \times \{0\}$$

■

Exemple 2 : Considérons la fonction polyédrale $f : I\!R^n \longrightarrow I\!R$ définie par

$$f(x) = \max_{1 \leq i \leq p} (a_i x + b_i)$$

Nous savons déjà [cfr proposition I.2.2.3] que $f''(x_0; d) = 0 \; \forall \; d$. Dès lors, pour $x_0^* \in \partial f(x_0)$, nous aurons que $f''(x_0, x_0^*; .)$ est la fonction indicatrice du cône normal $N(\partial f(x_0), x_0^*)$. Nous prouvons à présent que $\partial^2 f(x_0, x_0^*)$ se réduit au cône tangent $T(\partial f(x_0), x_0^*)$.
En effet,

$$\begin{aligned} f''(x_0, x_0^*; .) &= \delta(N(\partial f(x_0), x_0^*) \mid .) \\ &= \delta^*(T(\partial f(x_0), x_0^* \mid .) \qquad \text{[voir (7)]} \end{aligned}$$

et

$$\begin{aligned} \delta^*(\partial^2(f(x_0, x_0^*) \mid .) &= cl\sqrt{f''(x_0, x_0^*; .)} \\ &\leq \sqrt{f''(x_0, x_0^*; .)} \qquad \text{[cfr A.II.6]} \\ &\leq f''(x_0, x_0^*; .) \end{aligned}$$

Ainsi, nous obtenons grâce à la proposition A.VIII.1,

$$\partial^2 f(x_0, x_0^*) \subseteq T(\partial f(x_0), x_0^*)$$

L'égalité s'établit finalement au moyen de la proposition 2.4.
Quant au différentiel du second ordre f au point x_0, il n'est rien d'autre que le singleton $\{0\}$ [cfr proposition 2.1]. ■

<u>Exemple 3 :</u> Soient les fonctions $f_1, f_2 : I\!R^2 \to I\!R$ définies au moyen des deux égalités suivantes

$$\begin{aligned} f_1(x_1, x_2) &= x_1^2 + 2x_1 + 2x_2^2 \\ f_2(x_1, x_2) &= 2x_1^2 + x_2^2 + 2x_2 \end{aligned}$$

Nous posons $f = \max(f_1, f_2)$ et désirons calculer $\partial^2 f(x_0)$ au point $x_0 = (0,0)$.
Puisque f_1 et f_2 constituent des formes quadratiques, elles sont deux fois différentiables. Dès lors, la proposition 2.5 affirme que

$$\begin{aligned} & \partial^2 f_1(x_0) = E_1 = [\nabla^2 f_1(x_0)]^{1/2}(\overline{B}) \\ \text{et} \quad & \partial^2 f_2(x_0) = E_2 = [\nabla^2 f_2(x_0)]^{1/2}(\overline{B}) \end{aligned}$$

ou encore [cfr proposition I.2.4.2.1]

$$\partial^2 f_1(x_0) = \{(u,v) \mid (u,v)[\nabla^2 f_1(x_0)]^{-1} \begin{pmatrix} u \\ v \end{pmatrix} \leq 1\}$$

et

$$\partial^2 f_2(x_0) = \{(u,v) \mid (u,v)[\nabla^2 f_2(x_0)]^{-1} \begin{pmatrix} u \\ v \end{pmatrix} \leq 1\}$$

Ainsi, nous obtenons

$$\begin{aligned} \partial^2 f_1(x_0) &= \{(u,v) \mid \frac{u^2}{2} + \frac{v^2}{4} \leq 1\} \\ \partial^2 f_2(x_0) &= \{(u,v) \mid \frac{u^2}{4} + \frac{v^2}{2} \leq 1\} \end{aligned}$$

Par ailleurs, faisant appel à la proposition A.IV.7, nous établissons que

$$f'(x_0;d) = \max_{i \in I(x_0)} f_i'(x_0;d)$$

avec $I(x_0) = \{i \mid f_1(x_0) = f(x_0) = 0\} = \{1,2\}$.
Donc,

$$\begin{aligned} f'(x_0;d) &= \max\left(\nabla f_1(x_0)^T d, \nabla f_2(x_0)^T d\right) \\ &= \max\left((2,0)^T d, (0,2)^T d\right) \\ &= \max\left(2d_1, 2d_2\right) \\ &= 2\max\left(d_1, d_2\right) \end{aligned}$$

Nous avons aussi

$$f''(x_0;d) = \max_{i \in I(x_0;d)} f_i''(x_0;d)$$

où

$$\begin{aligned} I(x_0;d) &= \{i \in I(x_0) \mid f_i'(x_0;d) = f'(x_0;d)\} \\ &= \begin{cases} \{1\} & \text{si} \quad d_1 > d_2 \\ \{2\} & \text{si} \quad d_2 > d_1 \\ \{1,2\} & \text{si} \quad d_1 = d_2 \end{cases} \end{aligned}$$

Par conséquent, comme

$$\begin{aligned} f_1''(x_0;d) &= d^T[\nabla^2 f_1(x_0)]d \\ &= (d_1 d_2)\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \\ &= 2d_1^2 + 4d_2^2 \end{aligned}$$

et

$$\begin{aligned} f_2''(x_0;d) &= d^T[\nabla^2 f_2(x_0)]d \\ &= 4d_1^2 + 2d_2^2 \end{aligned}$$

Nous obtenons

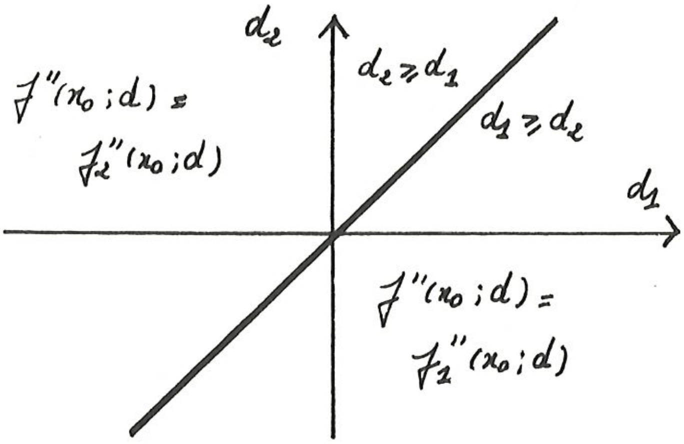


Figure 15

Considérant à présent la propriété A.IV.6.b.,

$$\begin{aligned}\partial f(x_0) &= Co\{\nabla f_1(x_0), \nabla f_2(x_0\} \\ &= Co\{(2,0),(0,2)\}\end{aligned}$$

ce qui nous oblige, pour calculer $\partial^2 f(x_0, x_0^*), x_0^* \in \partial f(x_0)$, à distinguer les trois possibilités suivantes

a) $x_0^* = (2,0)$:

Sous cette hypothèse,

$$\begin{aligned}d \in N(\partial f(x_0), x_0^*) &\Leftrightarrow\ < x_0^*, d >= f'(x_0,d) = \max(2d_1, 2d_2) \\ &\Leftrightarrow\ 2d_1 = \max(2d_1, 2d_2) \\ &\Leftrightarrow\ d_1 \geq d_2\end{aligned}$$

Dès lors,

$$f''(x_0, x_0^*; d) = \begin{cases} f''(x_0;d) = f_1''(x_0;d) = 2d_1^2 + 4d_2^2 & \text{si } d_1 \geq d_2 \\ +\infty & \text{sinon}\end{cases}$$

et

$$\begin{aligned}\partial^2 f(x_0, x_0^*) = \{x_0^{**}, (x_{01}, x_{02}) \mid x_{01}d_1 + x_{02}d_2 \leq \sqrt{2d_1^2 + 4d_2^2} \\ \forall\ (d_1, d_2) \text{ avec } d_1 \geq d_2\}\end{aligned}$$

Il est admis que cet ensemble n'est rien d'autre que l'ellipsoïde $\partial^2 f_1(x_0)$ translaté dans la direction $(-1, 1)$, c'est-à-dire

$$\partial^2 f(x_0, x_0^*) = \{z + \alpha(-1, 1) \mid \alpha \geq 0 \text{ et } \frac{z_1^2}{2} + \frac{z_2^2}{4} \leq 1 \ \ \forall\, z = (z_1, z_2)\}$$

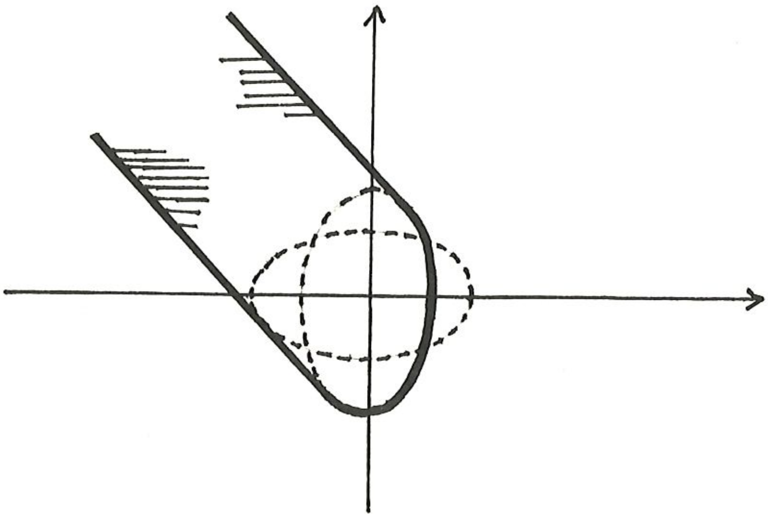


Figure 16

b) $x_0^* = (0, 2)$:

De façon tout à fait analogue au cas a), nous obtenons

$$d \in N(\partial f(x_0), x_0^*) \Leftrightarrow d_1 \leq d_2$$

$$f''(x_0, x_0^*; d) = \begin{cases} f''(x_0; d) = f_2''(x_0; d) = 4d_1^2 + 2d_2^2 & \text{si } d_1 \leq d_2 \\ +\infty & \text{sinon} \end{cases}$$

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} \mid < x_0^{**}, d >= \sqrt{4d_1^2 + 2d_2^2} \ \ \forall\, (d_1, d_2) = d \text{ avec } d_1 \leq d_2\}$$

ce qui donne géométriquement,

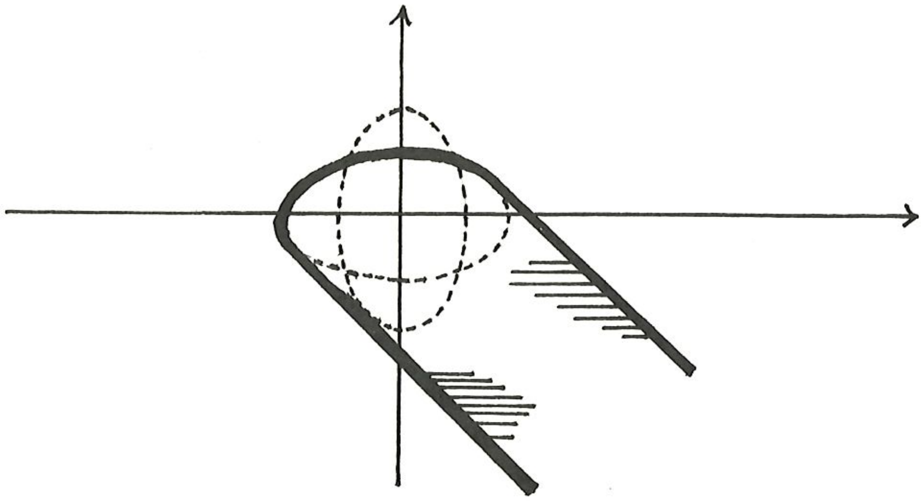


Figure 17

c) $x_0^* = (\alpha, 2-\alpha)$ avec $\alpha \in]0,2[$:

Ici,

$$\begin{aligned} d \in N(\partial f(x_0), x_0^*) &\Leftrightarrow \alpha d_1 + (2-\alpha) d_2 = \max(2d_1, 2d_2) \\ &\Leftrightarrow d_1 = d_2 \end{aligned}$$

Par conséquent,

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} | < x_0^{**}, d > \leq \sqrt{6 d_1^2} \ \forall \, (d_1, d_2) = d \text{ avec } d_2 = d_1\}$$

Cette condition sur $x_0^{**} = (x_{01}, x_{02})$ peut encore s'écrire

$$(x_{01} + x_{02}) d_1 \leq \sqrt{6} \, |d_1| \ \ \forall \ d_1$$

et ainsi

$$\partial^2 f(x_0, x_0^*) = \{x_0^{**} | -\sqrt{6} \leq x_{01} + x_{02} \leq \sqrt{6}\}$$

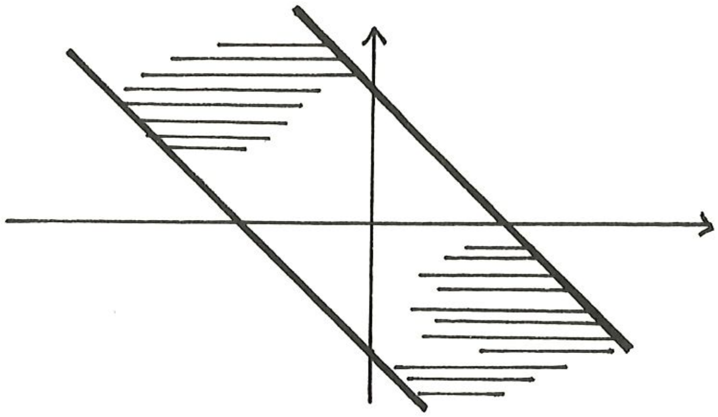


Figure 18

Nous considérons finalement l'intersection des différentiels seconds de f en (x_0, x_0^*) $\forall\ x_0^* \in \partial f(x_0)$ et obtenons donc pour $\partial^2 f(x_0)$

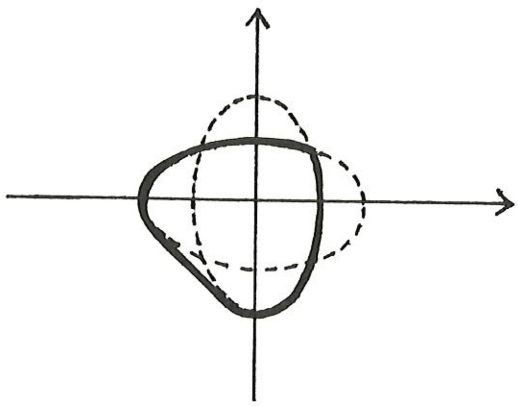


Figure 19

■

Exemple 4 : Considérons $f = ||\cdot||$ la norme euclidienne dans $\mathbb{R}^n$ et $x_0 = 0$. Nous obtenons immédiatement que

$$\begin{aligned}\partial f(x_0) &= \{x_0^* \in \mathbb{R} \mid\ ||x||\ \geq < x_0^*, x > \ \forall\ x \in \mathbb{R}^n\} \\ &= \{x_0^* \in \mathbb{R}^n \mid\ ||x_0^*||\ \leq 1\}\end{aligned}$$

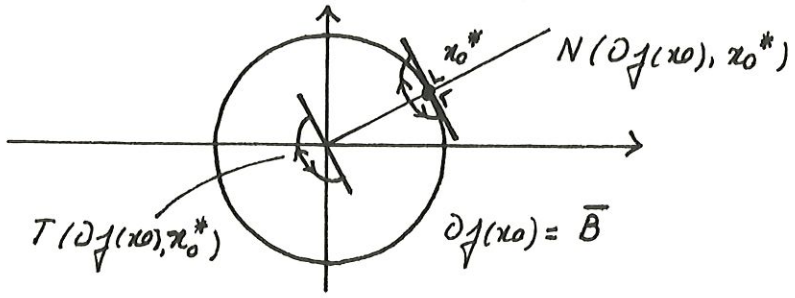


Figure 20

Dès lors, si $||x_0^*|| < 1$, nous pouvons affirmer [cfr proposition 1.1] que $\partial^2 f(x_0, x_0^*) = \mathbb{R}^n$. Par contre, si $||x_0^*|| = 1$, nous déduisons

$$\begin{aligned} f''(x_0; d) &= \lim_{\lambda \to 0^+} \frac{2}{\lambda}\left[\frac{f(x_0+\lambda d) - f(x_0)}{\lambda} - f'(x_0; d)\right] \\ &= \lim_{\lambda \to 0^+} \frac{2}{\lambda}\left[\frac{||\lambda d||}{\lambda} - \lim_{\lambda' \to 0^+} \frac{||\lambda' d||}{\lambda'}\right] \\ &= 0 \end{aligned}$$

et

$$f''(x_0, x_0^*; d) = \delta(\mathbb{R}_+\{x_0^*\}|d)$$

où

$$\mathbb{R}_+\{x_0^*\} = \{d \in \mathbb{R}^n \mid \exists \lambda \geq 0 \quad d = \lambda x_0^*\}$$

Ainsi,

$$\begin{aligned} \partial^2 f(x_0, x_0^*) &= \{x_0^{**} | < x_0^{**}, d > \leq 0 \ \ \forall\, d = \lambda x_0^*\} \\ &= T(\partial f(x_0), x_0^*) \end{aligned}$$

et donc

$$\partial^2 f(x_0) = \bigcap_{||x_0^*||=1} T(\partial f(x_0), x_0^*) = \{0\}$$

■

3 Résultats complémentaires et méthode de Newton

3.1 Différentiels seconds de deux fonctions comparables dans un voisinage de x_0

Soient deux fonctions convexes f et g de $I\!R^n$ dans $\overline{I\!R}$ telles que

$$\begin{cases} f(x_0) = g(x_0) \\ f(x) \leq g(x) \text{ dans un voisinage de } x_0 \end{cases}$$

Si f et g sont deux fois différentiables au point x_0, nous savons que

$$\nabla f(x_0) = \nabla g(x_0) \text{ et } < \nabla^2 f(x_0)d, d > \leq < \nabla^2 g(x_0)d, d > \ \forall\ d$$

Envisageons à présent le cas où ni f, ni g, ne sont différentiables en x_0. Nous avons alors

$$\begin{aligned} \partial f(x_0) &= \{x^* | f(x) \geq f(x_0) + < x^*, x - x_0 > \ \forall\ x \in I\!R^n\} \\ &\subseteq \partial g(x_0) = \{x^* | g(x) \geq g(x_0) + < x^*, x - x_0 > \ \forall\ x \in I\!R^n\} \end{aligned}$$

La propriété suivante compare les différentiels seconds de f et de g.

<u>**Proposition 3.1.1 :**</u>

$$\partial^2 f(x_0, x_0^*) \subseteq \partial^2 g(x_0, x_0^*) \ \forall\ x_0^* \in \partial f(x_0) \tag{13}$$

D'où,

$$\partial^2 f(x_0) \subseteq \bigcap_{x_0^* \in\, bd\partial f(x_0) \cap bd\partial g(x_0)} \partial^2 g(x_0, x_0^*) \tag{14}$$

en prenant comme convention que $\bigcap_{\emptyset}$ *représente l'espace* $I\!R^n$ *tout entier.*

<u>**Démonstration :**</u>

Soit x_0^* un point arbitraire de $\partial f(x_0)$. Puisque $\partial f(x_0) \subset \partial g(x_0)$, nous obtenons que

$$\begin{aligned} N(\partial g(x_0), x_0^*) &= \{d \mid \forall\ x^* \in \partial g(x_0) \ 0 \geq < d, x^* - x_0^* >\} \\ &\subseteq N(\partial f(x_0), x_0^*) \end{aligned}$$

Prenons d dans $N(\partial g(x_0), x_0^*)$ ce qui correspond à prendre la paire associée (x_0^*, d). En vertu de l'inclusion précédente, (x_0^*, d) constitue encore une paire associée pour f. De sorte que,

$$\begin{aligned} < x_0^*, d > &= f'(x_0; d) \\ &= g'(x_0; d) \end{aligned}$$

Ainsi, $\forall\ d \in N(\partial g(x_0), x_0^*)$

$$\begin{aligned} f''(x_0; d) &= \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} - f'(x_0; d) \right] \\ &\leq g''(x_0; d) = \lim_{\lambda \to 0^+} \frac{2}{\lambda} \left[\frac{g(x_0 + \lambda d) - f(x_0)}{\lambda} - f'(x_0; d) \right] \end{aligned}$$

D'où,

$$f''(x_0, x_0^*; d) \leq g''(x_0, x_0^*; d) \ \ \forall\ d$$

L'inclusion (13) est alors prouvée.

Le résultat (14) suit de la définition même de $\partial^2 f(x_0)$ en se rappelant que seuls les points frontières jouent un rôle significatif pour l'intersection. ■

3.2 Développements du second ordre

Nous fournissons maintenant des estimations du second ordre de $f(x_0 + \lambda d)$ en termes de $\partial^2 f(x_0, x_0^*)$ et $\partial^2 f(x_0)$.

Proposition 3.2.1 :

$\forall\ \lambda > 0$

a)

$$f(x_0 + \lambda d) \geq f(x_0) + \lambda < x_0^*, d > + \frac{\lambda^2}{2}[\delta^*(\partial^2 f(x_0, x_0^*)|d)]^2 + o(\lambda^2) \qquad (15)$$

$\forall\ x_0^* \in \partial f(x_0)$ tel que (x_0^*, d) forme une paire associée.

Nous avons l'égalité quand $f''(x_0, x_0^*; \cdot)$ est semi-continue inférieurement en d.

b)

$$f(x_0 + \lambda d) \geq f(x_0) + \lambda \delta^*(\partial f(x_0)|d) + \frac{\lambda^2}{2}[\delta^*(\partial^2 f(x_0)|d)]^2 + o(\lambda^2) \qquad (16)$$

L'égalité s'obtient lorsque $f''(x_0; \cdot)$ est convexe.

Démonstration :

Prouvons tout d'abord le résultat auxiliaire suivant

$$cl\sqrt{g(d)} = \sqrt{cl\ g(d)} \tag{17}$$

pour toute fonction g convexe, propre et positive.
En effet, faisant appel à la proposition A.II.7, il nous suffit de démontrer que

$$\begin{aligned}\sqrt{\liminf_{d' \to d} g(d')} &= \liminf_{d' \to d} \sqrt{g(d')} \\ &= \sup_{\varepsilon > 0} \inf_{|d'-d|<\varepsilon} \sqrt{g(d')}\end{aligned}$$

Dès lors, il nous reste à établir que

$$\inf_{|d'-d|\leq\varepsilon} \sqrt{g(d')} = \sqrt{\inf_{|d'-d|\leq\varepsilon} g(d')}$$

(Un raisonnement similaire nous autorisera à permuter le supremum et la racine).
Nous avons premièrement

$$\sqrt{g(d')} \geq \sqrt{\inf_{d' \in \bar{B}(d,\varepsilon)} g(d')} \quad \forall\ d' \in \bar{B}(d,\varepsilon)$$

et donc

$$\inf_{d' \in \bar{B}(d,\varepsilon)} \sqrt{g(d')} \geq \sqrt{\inf_{d' \in \bar{B}(d,\varepsilon)} g(d')}$$

Ensuite, montrons que

$$\forall\ \beta > 0\ \ \exists\ d' \in \bar{B}(d,\varepsilon)\ \ \sqrt{g(d')} < \sqrt{\inf_{d' \in \bar{B}(d,\varepsilon)} g(d')} + \beta$$

Comme l'infimum est le plus grand des minorants, nous obtenons

$$\forall\ \delta > 0\ \ \exists d'' \in \bar{B}(d,\varepsilon)\ \ g(d'') < \inf_{d' \in \bar{B}(d,\varepsilon)} g(d') + \delta$$

ce qui implique que

$$\begin{aligned}\sqrt{g(d'')} &< \sqrt{\inf_{d' \in \bar{B}(d,\varepsilon)} g(d') + \delta} \\ &\leq \sqrt{\inf_{d' \in \bar{B}(d,\varepsilon)} g(d')} + \sqrt{\delta}\end{aligned}$$

Ceci clôture la preuve de l'égalité (17). Venons-en à l'énoncé même.

a) Nous savons déjà, par définition de la dérivée directionnelle du second ordre de f en x_0, que

$$f(x_0 + \lambda d) = f(x_0) + \lambda f'(x_0; d) + \frac{\lambda^2}{2} f''(x_0; d) + o(\lambda^2) \quad \forall\, \lambda > 0 \tag{18}$$

Puisque (x_0^*, d) constitue une paire associée

$$f'(x_0; d) = < x_0^*, d >$$

et

$$f''(x_0; d) = f''(x_0, x_0^*; d)$$

Or,

$$\begin{aligned} [\delta^*(\partial^2 f(x_0, x_0^*) \mid d)]^2 &= \left[cl \sqrt{f''(x_0, x_0^*; d)} \right]^2 \\ &= \left[\sqrt{cl\ f''(x_0; x_0; d)} \right]^2 \qquad \text{[cfr (17)]} \\ &= cl\ f''(x_0, x_0^*; d) \end{aligned}$$

donc,

$$\begin{aligned} f''(x_0; d) &= f''(x_0, x_0^*; d) \\ &\geq cl\ f''(x_0, x_0^*; d) \qquad \text{[cfr proposition A.II.6]} \\ &= [\delta^*(\partial^2 f(x_0, x_0^*) \mid d)]^2 \end{aligned}$$

De cette façon, l'inégalité (15) est obtenue.
Si l'on suppose en outre que $f''(x_0, x_0^*; d)$ est semi-continue inférieurement, alors, par la propriété A.II.5, $f''(x_0, x_0^*; d)$ est fermée et nous pouvons écrire

$$\begin{aligned} f''(x_0; d) = f''(x_0, x_0^*; d) &= cl\ f''(x_0, x_0^*; d) \\ &= [\delta^*(\partial^2 f(x_0, x_0^*) \mid d)]^2 \end{aligned}$$

Ainsi, l'inégalité fait place à une égalité dans (15).

b) Puisque $f'(x_0; d) = \delta^*(\partial f(x_0) \mid d)$ [cfr A.IV.5], il nous suffit de montrer, relativement à (18), que

$$f''(x_0; d) \geq [\delta^*(\partial^2 f(x_0)|d)]^2 = \left[\sup_{x \in \partial^2 f(x_0)} < x, d > \right]^2$$

l'égalité remplaçant l'inégalité dans le cas où $f''(x_0;.)$ est convexe.
Nous savons déjà que

$$\forall\, x \in \partial^2 f(x_0) \quad < x, d > \leq \sqrt{f''(x_0; d)} \ \forall\, d$$

Donc,

$$\sup_{x\in\partial^2 f(x_0)} \quad < x, d > \leq \sqrt{f''(x_0; d)} \ \forall\, d$$

comme $0 \in \partial^2 f(x_0)$, nous obtenons

$$0 \leq \left[\sup_{x\in\partial^2 f(x_0)} \quad < x, d >\right]^2 \leq f''(x_0; d)$$

De plus, si $f''(x_0;.)$ est convexe

$$\begin{aligned}
[\delta^*(\partial^* f(x_0) \mid d)]^2 &= \left[(\sqrt{f''(x_0; d)})^{**}\right]^2 \\
&= \left[cl\ \sqrt{f''(x_0; d)}\right]^2 && \text{[cfr A.III.3]} \\
&= \left[\sqrt{cl\ f''(x_0; d)}\right]^2 && \text{[cfr(17)]} \\
&= cl\ f''(x_0; d)
\end{aligned}$$

Puisque $f''(x_0; d)$ est une fonction convexe de domaine $I\!R^n$ et à valeurs réelles, elle est continue, donc semi-continue inférieurement. En se référant alors à la proposition A.II.5, la thèse suit. ■

Au moyen de cette proposition, nous pouvons généraliser la méthode de Newton décrite dans le chapitre d'introduction en réécrivant le problème (P) sous la forme

$$(Q_1) \equiv \begin{cases} \min \quad f'(x_0; d) + \frac{1}{2}[\delta^*(\partial^2 f(x_0) \mid d)]^2 \\ \text{sous} \quad d \in I\!R^n \end{cases}$$

A nouveau, nous prenons $x_0 \in I\!R^n$ comme point initial et définissons le point de l'itération suivante par $x_1 = x_0 + d_0$. Le problème (Q_1) présente l'avantage d'être convexe.

Nous pouvons envisager une interprétation géométrique dans le cas où f est différentiable. (Q_1) devient alors :

$$(Q_1) \quad \equiv \quad \begin{cases} \min & \nabla f(x_0)^T d + \frac{1}{2}[\delta^*(\partial^2 f(x_0) \mid d)]^2 \\ \text{sous} & d \in I\!R^n \end{cases}$$

$$\stackrel{\text{not}}{\equiv} \quad \begin{cases} \min & F(d) \\ \text{sous} & d \in I\!R^n \end{cases}$$

Considérons une solution d_0^*.
La condition nécessaire et suffisante d'optimalité du problème convexe (Q) [cfr A.VI.2] s'écrit ici

$$0 \in \partial F(d_0^*)$$

c'est-à-dire, en vertu des propriétés élémentaires du sous-différentiel [cfr A.IV.6]

$$0 \in \nabla f(x_0) + \delta^*(\partial^2 f(x_0) \mid d_0^*) \;\; \partial\delta^*(\partial^2 f(x_0) \mid .)(d_0^*)$$

Nous considérons que $\nabla f(x_0) \neq 0$ (sans quoi x_0 serait minimum puisque f est convexe). Dès lors, $d_0^* \neq 0$ et

$$\delta^*(\partial^2 f(x_0) \mid d_0^*) = \max_{x \in \partial^2 f(x_0)} \;\; < x, d_0^* > \; > 0$$

lorsque l'on suppose $\partial^2 f(x_0)$ non réduit à l'origine.
D'où,

$$\frac{-\nabla f(x_0)}{\delta^*(\partial^2 f(x_0) \mid d_0^*)} \in \partial\delta^*(\partial^2 f(x_0) \mid .)(d_0^*)$$

Nous appliquons maintenant le théorème A.IV.8 à la fonction indicatrice $\delta(\partial^2 f(x_0) \mid .)$ qui est bien convexe, propre, semi-continue inférieurement (puisque $\partial^2 f(x_0)$ est un convexe fermé).
Nous obtenons ainsi,

$$d_0^* \in \partial\delta(\partial^2 f(x_0) \mid .) \left(\frac{-\nabla f(x_0)}{\delta^*(\partial^2 f(x_0) \mid d_0^*)} \right)$$

où

$$\frac{-\nabla f(x_0)}{\delta^*(\partial^2 f(x_0) \mid d_0^*)} \in \operatorname{dom} \delta(\partial^2 f(x_0) \mid .) = \partial^2 f(x_0)$$

Par conséquent,

$$d_0^* \in N \left(\partial^2 f(x_0), \frac{-\nabla f(x_0)}{\delta^*(\partial^2 f(x_0) \mid d_0^*)} \right)$$

Comme $d_0^* \neq 0$, nous avons que

$$\frac{-\nabla f(x_0)}{\delta^*(\partial^2 f(x_0) \mid d_0^*)} \in bd\, \partial f(x_0)$$

Nous représentons ci-dessous l'ensemble des directions possibles pour d_0^*.

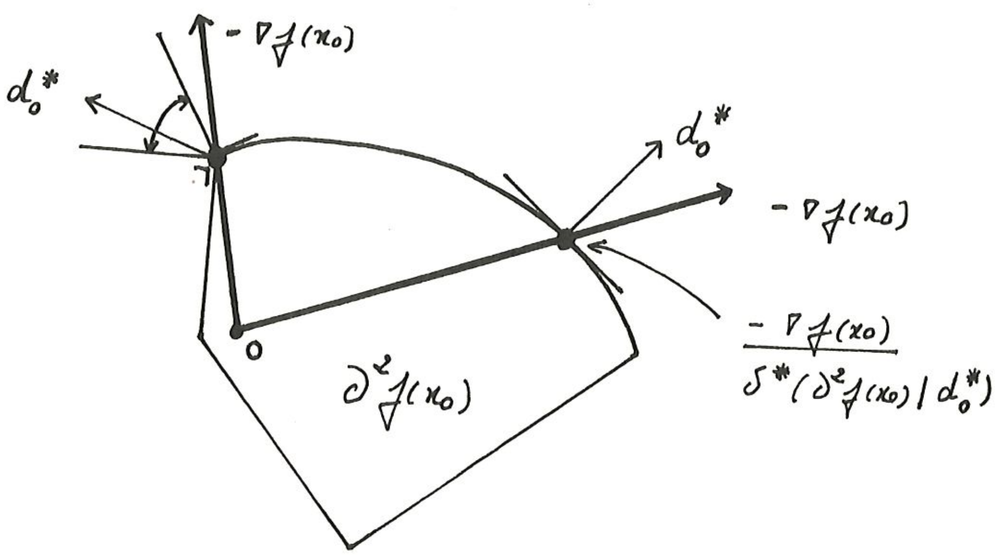


Figure 21

CHAPITRE III

REGLES DE CALCUL

Nous désirons, à présent, trouver les expressions des différentiels du second ordre d'une fonction f construite à partir d'autres fonctions f_i dont les propriétés sont mieux connues. Il s'agira donc d'exprimer $\partial^2 f$ en terme des $\partial^2 f_i$ plus faciles à calculer.
Pour ce faire, nous nous limiterons à trois opérations préservant la convexité, à savoir :

1 La multiplication d'une fonction par un nombre réel positif (situé à gauche de la fonction)

Rappelons que la multiplication à gauche par le scalaire $\alpha \geq 0$, produit la nouvelle fonction $\alpha f : x \to (\alpha f)x = \alpha f(x)$.

Proposition 1.1 :

$\forall\, \alpha \geq 0$, *nous avons que*

$$\partial^2(\alpha f)(x_0, x_0^*) = \sqrt{\alpha}\ \partial^2 f(x_0, x_0^*)\ \ \forall\ x_0^* \in \partial f(x_0)$$

et

$$\partial^2(\alpha f)(x_0) = \sqrt{\alpha}\ \partial^2 f(x_0)$$

Démonstration :

Constatons tout d'abord que

$$\begin{aligned}(\alpha f)'(x_0; d) &= \lim_{\lambda \to 0^+} \frac{(\alpha f)(x_0 + \lambda d) - (\alpha f)(x_0)}{\lambda} \quad \forall\ d \in I\!R^n \\ &= \alpha f'(x_0; d)\end{aligned}$$

et

$$\begin{aligned}(\alpha f)''(x_0;d) &= \lim_{\lambda\to 0^+} \frac{1}{\lambda}[(\alpha f)'(x_0+\lambda d;d)-(\alpha f)'(x_0;d)] \\ &= \alpha f''(x_0;d)\end{aligned}$$

Dès lors, nous obtenons que

$$\begin{aligned}\partial^2(\alpha f)(x_0,x_0^*) &= \{x_0^{**}|<x_0^{**},d>\leq \sqrt{(\alpha f)''(x_0,x_0^*;d)}\ \ \forall\ d\} \\ &= \{x_0^{**}|<x_0^{**},d>\leq \sqrt{(\alpha f)''(x_0;d)}\ \ \forall\ d\in N(\partial f(x_0),x_0^*)\} \\ &= \{x_0^{**}|<\frac{x_0^{**}}{\sqrt{\alpha}},d>\leq \sqrt{f''(x_0;d)}\ \ \forall\ d\in N(\partial f(x_0),x_0^*)\} \\ &= \{\sqrt{\alpha}z^{**}|\ z^{**}\in\partial^2 f(x_0,x_0^*)\} \\ &= \sqrt{\alpha}\ \ \partial^2 f(x_0,x_0^*)\end{aligned}$$

D'autre part,

$$\begin{aligned}\partial^2(\alpha f)(x_0) &= \{x_0^{**}|<x_0^{**},d>\leq \sqrt{(\alpha f)''(x_0;d)}\ \ \forall\ d\} \\ &= \{\sqrt{\alpha}z^{**}|\ z^{**}\in\partial^2 f(x_0)\} \\ &= \sqrt{\alpha}\ \ \partial^2 f(x_0)\end{aligned}$$

■

2 La somme de fonctions

Soient deux fonctions convexes f_1 et f_2, définies sur $I\!R^n$ et à valeurs dans $I\!R$ pour lesquelles la dérivée directionnelle du second ordre existe et est finie pour tout x_0 et tout d. Considérons la somme

$$f_1+f_2 : x \to (f_1+f_2)(x) = f_1(x)+f_2(x)$$

Nous savons que

$$(f_1+f_2)'(x_0;d) = f_1'(x_0;d)+f_2'(x_0;d)\quad \forall\ d \tag{19}$$

et

$$\partial(f_1+f_2)(x_0) = \partial f_1(x_0)+\partial f_2(x_0) \tag{20}$$

Proposition 2.1 :

Si

$$x_0^* = x_0^{*1} + x_0^{*2} \in \partial(f_1 + f_2)(x_0) \;\forall\; x_0^{*1} \in \partial f_1(x_0) \textit{ et } \forall\; x_0^{*2} \in \partial f_2(x_0)$$

Alors

a) $N(\partial(f_1 + f_2)(x_0), x_0^*) = N(\partial f_1(x_0), x_0^{*1}) \cap N(\partial f_2(x_0), x_0^{*2})$

b) $(f_1 + f_2)''(x_0; d) = f_1''(x_0; d) + f_2''(x_0; d)$

de sorte que $f_1 + f_2 = f$ constitue bien une fonction convexe, définie sur $\mathbb{R}^n$ et telle que $f''(x_0; d)$ existe et est finie.

Démonstration :

a) $d \in N(\partial(f_1 + f_2)(x_0), x_0^*)$

$\Leftrightarrow (x_0^*, d)$ forme une paire associée pour $f_1 + f_2$

$$\begin{aligned} \Leftrightarrow x_0^* \in \partial(f_1 + f_2)(x_0) \text{ et } < x_0^*, d > &= (f_1 + f_2)'(x_0; d) \\ &= f_1'(x_0; d) + f_2'(x_0; d) \qquad \text{[cfr(19)]} \end{aligned}$$

Comme $x_0^{*i} \in \partial f_i(x_0) \quad i = 1, 2$ nous avons aussi

$$< x_0^{*i}, d > \leq f_i'(x_0; d) \qquad \text{[voir proposition A.IV.3]}$$

Dès lors,

$$< x_0^{*i}, d > = f_i'(x_0; d)$$

et ainsi $d \in N(\partial f_1(x_0), x_0^{*1})$ et $d \in N(\partial f_2(x_0), x_0^{*2})$.
L'autre inclusion est tout aussi immédiate, en vertu de (19) et de (20).

b) Ce résultat s'obtient aisément au moyen de la définition de $(f_1, f_2)''(x_0; d)$ [cfr(2)] et de l'égalité (19). ∎

Proposition 2.2 :

Sous les mêmes hypothèses que la propriété 2.1, nous avons

$$\begin{aligned} \frac{1}{\sqrt{2}}\{\partial^2 f_1(x_0, x_0^{*1}) + \partial^2 f_2(x_0, x_0^{*2})\} &\subseteq \partial^2(f_1 + f_2)(x_0, x_0^*) \qquad (21) \\ &\subseteq \sqrt{2}\, Co\, \{\partial^2 f_1(x_0, x_0^{*1}) \cup \partial^2 f_2(x_0, x_0^{*2})\} \end{aligned}$$

et

$$\frac{1}{\sqrt{2}}\{\partial^2 f_1(x_0)+\partial^2 f_2(x_0)\}\subseteq \partial^2(f_1+f_2)(x_0) \tag{22}$$

<u>Démonstration :</u>

Des résultats de la proposition 2.1, nous pouvons déduire que

$$(f_1+f_2)''(x_0,x_0^*;d)=f_1''(x_0,x_0^{*1};d)+f_2''(x_0,x_0^{*2};d)\ \ \forall\ d \tag{23}$$

De plus, nous avons

$$\begin{aligned}\partial^2(f_1+f_2)(x_0,x_0^*) &= \{x_0^{**}|<x_0^{**},d>\leq \sqrt{(f_1+f_2)''(x_0,x_0^*;d)}\ \ \forall\ d\}\\ &= \partial\sqrt{(f_1+f_2)''(x_0,x_0^*;.)}(0)\\ &= \partial\sqrt{f_1''(x_0,x_0^{*1};.)+f_2''(x_0,x_0^{*2};.)}(0)\end{aligned}$$

en vertu de l'égalité (23).
Par ailleurs, nous utilisons l'inégalité générale

$$\sqrt{a+b}\leq\sqrt{2}\ \ \text{Max}\{\sqrt{a},\sqrt{b}\}\ \ \forall\ a\geq 0, b\geq 0$$

de sorte que

$$\begin{aligned}&\partial\sqrt{f_1''(x_0,x_0^{*1};.)+f_2''(x_0,x_0^{*2};.)}(0)\\ &\quad= \{x_0^{**}|<x_0^{**},d>\leq\sqrt{f_1''(x_0,x_0^{*1};.)+f_2''(x_0,x_0^{*2};.)}\ \ \forall\ d\}\\ &\quad\subseteq \{x_0^{**}|<x_0^{**},d>\leq\sqrt{2}\ \ \text{Max}\{\sqrt{f_1''(x_0,x_0^{*1};.)},\sqrt{f_2''(x_0,x_0^{*2};.)}\}\ \forall\ d\}\\ &\quad= \sqrt{2}\ \ \partial[\text{Max}\{\sqrt{f_1''(x_0,x_0^{*1};.)},\sqrt{f_2''(x_0,x_0^{*2};.)}\}](0)\\ &\quad= \sqrt{2}\,Co\ [\partial\sqrt{f_1''(x_0,x_0^{*1};.)}(0)\cup\partial\sqrt{f_2''(x_0,x_0^{*2};.)}(0)]\end{aligned}$$

la dernière ligne pouvant s'établir au moyen de la proposition A.IV.6.b.
Nous obtenons par conséquent

$$\begin{aligned}\partial^2(f_1+f_2)(x_0,x_0^*) &\subseteq \sqrt{2}\,Co[\partial\sqrt{f_1''(x_0,x_0^{*1};.)}(0)\cup\partial\sqrt{f_2''(x_0,x_0^{*2};.)}]\\ &= \sqrt{2}\,Co[\partial^2 f_1(x_0,x_0^{*1})\cup\partial^2 f_2(x_0,x_0^{*2})]\end{aligned}$$

Concernant l'estimation inférieure de $\partial^2(f_1+f_2)(x_0,x_0^*)$, nous nous servons cette fois de l'inégalité générale

$$\frac{1}{\sqrt{2}}(\sqrt{a}+\sqrt{b})\leq\sqrt{a+b}\ \ \ \forall\ a\geq 0, b\geq 0$$

provenant du caractère concave de la fonction racine carrée.

Dès lors,

$$\begin{aligned}
\partial^2(f_1+f_2)(x_0,x_0^*) &= \partial\sqrt{(f_1+f_2)''(x_0,x_0^*;.)}(0) \\
&= \partial\sqrt{f_1''(x_0,x_0^{*1};.)+f_2''(x_0,x_0^{*2};.)}(0) \qquad \text{[cfr (23)]} \\
&\supseteq \frac{1}{\sqrt{2}}\,\partial\,(\sqrt{f_1''(x_0,x_0^{*1};.)}+\sqrt{f_2''(x_0,x_0^{*2};.)}](0) \\
&= \frac{1}{\sqrt{2}}[\partial\sqrt{f_1''(x_0,x_0^{*1};.)}(0)+\partial\sqrt{f_2''(x_0,x_0^{*2};.)}(0)] \\
&= \frac{1}{\sqrt{2}}[\partial^2 f_1(x_0,x_0^{*1})+\partial^2 f_2(x_0,x_0^{*2})]
\end{aligned}$$

Finalement, nous prouvons l'inclusion (22).
Soit

$$z^{**} \in \frac{1}{\sqrt{2}}\{\partial^2 f_1(x_0)+\partial^2 f_2(x_0)\}$$

c'est-à-dire

$$z^{**} \in \frac{1}{\sqrt{2}}\{\partial^2 f_1(x_0,x_0^{*1})+\partial^2 f_2(x_0,x_0^{*2})\} \quad \forall\, x_0^{*1}\in\partial f_1(x_0), x_0^{*2}\in\partial f_2(x_0)$$

Il suit ainsi de la première inclusion de (21) que

$$z^{**}\in\partial^2(f_1+f_2)(x_0,x_0^*) \;\; \forall\, x_0^*\in\partial(f_1+f_2)(x_0)$$

donc,

$$z^{**}\in\partial^2(f_1+f_2)(x_0)$$

■

Nous donnons à présent des estimations inférieures de $\partial^2(f_1+f_2)(x_0,x_0^*)$ et $\partial^2(f_1+f_2)(x_0)$ du type enveloppe convexe.

Proposition 2.3 :

Considérons

$$x_0^* = x_0^{*1}+x_0^{*2}\in\partial(f_1+f_2)(x_0) \;\; \forall\, x_0^{*1}\in\partial f_1(x_0), x_0^{*2}\in\partial f_2(x_0)$$

Nous avons alors

$$Co\{\partial^2 f_1(x_0,x_0^{*1})\cup\partial^2 f_2(x_0,x_0^{*2})\}\subseteq\partial^2(f_1+f_2)(x_0,x_0^*)$$

et

$$Co\{\partial^2 f_1(x_0) \cup \partial^2 f_2(x_0)\} \subseteq \partial^2(f_1+f_2)(x_0)$$

<u>**Démonstration :**</u>

Au moyen de l'inégalité

$$\sqrt{f_1''(x_0, x_0^{*1};.)} \leq \sqrt{(f_1+f_2)''(x_0, x_0^*;.)}$$

nous déduisons

$$\begin{aligned}
\partial^2 f_1(x_0, x_0^{*1}) &= \partial\sqrt{f_1''(x_0, x_0^{*1};.)}(0) \\
&= \{x_0^{**} | < x_0^{**}, d > \leq \sqrt{f_1''(x_0, x_0^{*1}; d)} \ \ \forall\ d\} \\
&\subseteq \partial\sqrt{(f_1+f_2)''(x_0, x_0^*;.)}(0) \\
&= \partial^2(f_1+f_2)(x_0, x_0^*)
\end{aligned}$$

De façon similaire, nous obtenons

$$\partial^2 f_2(x_0, x_0^{*2}) \subseteq \partial^2(f_1+f_2)(x_0, x_0^*)$$

Comme le différentiel du second ordre de (f_1+f_2) en (x_0, x_0^*) est un ensemble convexe, nous retrouvons la première inclusion annoncée.
Quant à la seconde, elle s'obtient en considérant l'inégalité

$$f_1''(x_0;.) \leq (f_1+f_2)''(x_0;.)$$

de laquelle nous tirons

$$\begin{aligned}
\partial^2 f_1(x_0) &= \partial\sqrt{f_1''(x_0;.)}(0) \\
&= \{x_0^{**} | < x_0^{**}, d > \leq \sqrt{f_1''(x_0, d)} \ \ \forall\ d\} \\
&\subseteq \partial\sqrt{(f_1+f_2)''(x_0;.)}(0) \\
&= \partial^2(f_1+f_2)(x_0)
\end{aligned}$$

De manière analogue, nous pouvons montrer que

$$\partial^2 f_2(x_0) \subseteq \partial^2(f_1+f_2)(x_0)$$

Ainsi, $\partial^2(f_1+f_2)(x_0)$, ensemble convexe, contient bien l'enveloppe convexe de $U.\partial^2 f_i(x_0) \quad i=1,2.$ ∎

Proposition 2.4 :

Si $f_1''(x_0;.)$ et $f_2''(x_0;.)$ sont des fonctions convexes, nous avons

$$\partial^2(f_1+f_2)(x_0) \subseteq \partial^2 f_1(x_0) + \partial^2 f_2(x_0)$$
$$\partial^2(f_1+f_2)(x_0) \subseteq \sqrt{2}\, Co\{\partial^2 f_1(x_0) \cup \partial^2 f_2(x_0)\}$$

Démonstration :

En effet,

$$\begin{aligned}
\partial^2(f_1+f_2)(x_0) &= \partial\sqrt{(f_1+f_2)''(x_0;.)}(0)\\
&= \partial\sqrt{f_1''(x_0;.)+f_2''(x_0;.)}(0)\\
&\subseteq \partial[\sqrt{f_1''(x_0;.)}+\sqrt{f_2''(x_0;.)}](0)\\
&= \partial\sqrt{f_1''(x_0;.)}(0)+\partial\sqrt{f_2''(x_0;.)}(0)\\
&= \partial^2 f_1(x_0)+\partial^2 f_2(x_0)
\end{aligned}$$

Par ailleurs, au moyen de l'inégalité générale,

$$\sqrt{a+b} \leq \sqrt{2}\ \max(\sqrt{a},\sqrt{b})\ \ \forall\ a \geq 0, b \geq 0$$

nous obtenons

$$\begin{aligned}
\partial^2(f_1+f_2)(x_0) &= \partial\sqrt{f_1''(x_0;.)+f_2''(x_0;.)}(0)\\
&\subseteq \sqrt{2}\,\partial\,[\mathrm{Max}\{\sqrt{f_1''(x_0;.)},\sqrt{f_2''(x_0;.)}\}](0)\\
&= \sqrt{2}\, Co\,[\partial\sqrt{f_1''(x_0;.)}(0)\cup\partial\sqrt{f_2''(x_0;.)}(0)]
\end{aligned}$$

en nous référant à la proposition A.IV.6.b.
La thèse s'en suit. ∎

L'exemple que nous considérons maintenant nous montre l'importance de l'hypothèse de convexité des $f_i''(x_0;d)\quad i=1,2$ dans la proposition précédente.

Exemple 5 :

Soient $f_1, f_2 : I\!R^2 \to I\!R$ définies par

$$f_1(x) = \mathrm{Max}\{\frac{1}{2}x_1^2+\frac{1}{2}x_2^2+x_1, 0\}$$
$$f_2(x) = \mathrm{Max}\{\frac{1}{2}x_1^2+\frac{1}{2}x_2^2-x_1, 0\}$$

Pour éviter les calculs inutiles, reprenons les résultats de l'exemple 1 vu précédemment et adaptons les à notre contexte au moyen des règles de calcul étudiées ci-dessus.
Nous avons considéré, dans l'exemple 1, la fonction

$$\begin{array}{rlcl} f: & \mathbb{R}^2 & \rightarrow & R \\ & x=(x_1,x_2) & \rightsquigarrow & f(x)=\max\{0, x_1^2+x_2^2+2x_1\} \end{array}$$

Dès lors, nous obtenons, pour $x_0=(0,0)$,

$$f_1(x)=\frac{1}{2}f(x)=(\frac{1}{2}f)(x)$$

$$\begin{aligned}
f_1'(x_0;d) &= (\frac{1}{2}f)'(x_0;d)=\frac{1}{2}f'(x_0;d) \\
&= \frac{1}{2}\begin{cases} 0 & \text{si } d_1\leq 0 \\ 2d_1 & \text{si } d_1>0 \end{cases} \\
&= \begin{cases} 0 & \text{si } d_1\leq 0 \\ d_1 & \text{si } d_1>0 \end{cases} \\
&= \max(0,d_1) \\
f_1''(x_0;d) &= (\frac{1}{2}f)''(x_0;d)=\frac{1}{2}f''(x_0;d) \\
&= \frac{1}{2}\begin{cases} 2d_1^2+2d_2^2 & \text{si } d_1\geq 0 \\ 0 & \text{si } d_1<0 \end{cases} \\
&= \begin{cases} d_1^2+d_2^2 & \text{si } d_1\geq 0 \\ 0 & \text{si } d_2<0 \end{cases} \\
\partial^2 f_1(x_0) &= \partial^2(\frac{1}{2}f)(x_0)=\frac{1}{\sqrt{2}}\,\partial^2 f(x_0) \\
&= \frac{1}{\sqrt{2}}\{[0,\sqrt{2}]\times\{0\}\} \\
&= [0,1]\times\{0\}
\end{aligned}$$

Calculons à présent le différentiel second de f_2 en x_0.

$$f_2(x)=\frac{1}{2}\max\{x_1^2+x_2^2-2x_1,0\}=\frac{1}{2}F(x)$$

par notation.
Dès lors,

$$F(x) = \begin{cases} 0 & \text{sur le disque } (x_1-1)^2 + x_2^2 - 1 \leq 0 \\ x_1^2 + x_2^2 - 2x_1 & \text{ailleurs} \end{cases}$$

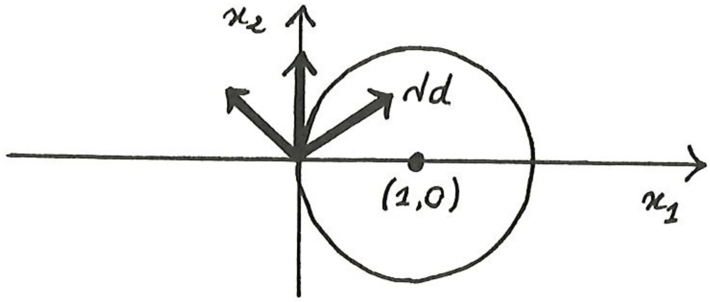


Figure 22

Nous avons

$$\begin{aligned} F'(x_0; d) &= \lim_{\lambda \to 0+} \frac{F(x_0 + \lambda d) - F(x_0)}{\lambda} \\ &= \lim_{\lambda \to 0+} \frac{F(\lambda d)}{\lambda} \end{aligned}$$

a) Si $d_1 > 0$:

Pour λ assez petit, nous avons que $(\lambda d_1, \lambda d_2)$ appartient au disque $\mathcal{D}$ dont l'équation est reprise ci-dessus. Ainsi,

$$F(\lambda d) = 0 \text{ et}$$
$$F'(x_0; d) = 0$$

b) Si $d_1 \leq 0$:

Cette fois, $(\lambda d_1, \lambda d_2)$ ne se trouve plus dans le disque $\mathcal{D}$ et nous obtenons

$$\begin{aligned} F'(x_0; d) &= \lim_{\lambda \to 0+} \frac{\lambda^2 d_1^2 + \lambda^2 d_2^2 - 2\lambda d_1}{\lambda} \\ &= -2d_1 \end{aligned}$$

Dès lors,

$$\begin{aligned} f_2'(x_0; d) &= (\frac{1}{2}F)'(x_0; d) = \frac{1}{2}F'(x_0; d) \\ &= \frac{1}{2}\max(0, -2d_1) \\ &= \max(0, -d_1) \end{aligned}$$

De même,

$$f_2''(x_0;d) = (\frac{1}{2}F)''(x_0;d) = \frac{1}{2}F''(x_0;d)$$

où

$$F''(x_0;d) = \lim_{\lambda\to 0^+} \frac{2}{\lambda}\left[\frac{F(x_0+\lambda d)-F(x_0)}{\lambda} - F'(x_0;d)\right]$$

a) Si $d_1 > 0$:

$$F''(x_0;d) = \lim_{\lambda\to 0^+} \frac{2}{\lambda}[0-0] = 0$$

b) Si $d_1 \leq 0$:

$$\begin{aligned} F''(x_0;d) &= \lim_{\lambda\to 0^+} \frac{2}{\lambda}\left[\frac{\lambda^2 d_1^2 + \lambda^2 d_2^2 - 2\lambda d_1}{\lambda} - (-2d_1)\right] \\ &= \lim_{\lambda\to 0^+} 2[d_1^2 + d_2^2] \\ &= 2d_1^2 + 2d_2^2 \end{aligned}$$

Par conséquent,

$$f_2''(x_0;d) = \begin{cases} d_1^2 + d_2^2 & \text{si } d_1 \leq 0 \\ 0 & \text{si } d_1 > 0 \end{cases}$$

et

$$\begin{aligned} \partial^2 f_2(x_0) &= \{x_0^* | < x_0^*, d > \leq \sqrt{f_2''(x_0;d)} \ \forall\ d\} \\ &= \{x_0^* | < x_0^*, d > \leq \begin{Bmatrix} \sqrt{d_1^2+d_2^2} & \text{si } d_1 \leq 0 \\ 0 & \text{si } d_1 > 0 \end{Bmatrix} \ \forall\ d_2\} \end{aligned}$$

Réécrivons ces conditions pour $x_0^* = (x_{01}^*, x_{02}^*)$

$$\begin{cases} x_{01}^* d_1 + x_{02}^* d_2 \leq \sqrt{d_1^2+d_2^2} & \text{si } d_1 \leq 0 \\ x_{01}^* d_1 + x_{02}^* d_2 \leq 0 & \text{si } d_1 > 0 \end{cases}$$

La deuxième contrainte nous donne

$$x_{01}^* \leq 0 \text{ et } x_{02}^* = 0$$

et la première nous permet d'obtenir finalement comme ensemble de solutions

$$\partial^2 f_2(x_0) = [-1, 0] \times \{0\}$$

Il nous reste ainsi à trouver le différentiel du second ordre de $(f_1 + f_2)$ au point x_0. Calculons donc

$$\begin{aligned}
(f_1 + f_2)''(x_0; d) &= f_1''(x_0; d) + f_2''(x_0; d) \\
&= \begin{cases} d_1^2 + d_2^2 \\ 0 \\ d_2^2 \end{cases} + \begin{cases} 0 & \text{si } d_1 > 0 \\ d_1^2 + d_2^2 & \text{si } d_1 < 0 \\ d_2^2 & \text{si } d_1 = 0 \end{cases} \\
&= \begin{cases} 2d_2^2 & \text{si } d_2 = 0 \\ d_1^2 + d_2^2 & \text{sinon} \end{cases}
\end{aligned}$$

de sorte que

$$\partial^2 (f_1 + f_2)(x_0) = \{x_0^* | < x_0^*, d > \leq \left\{ \begin{array}{ll} \sqrt{2d_2^2} & \text{si } d_1 = 0 \\ \sqrt{d_1^2 + d_2^2} & \text{si } d_1 \neq 0 \end{array} \right\} \ \forall\ d_2\}$$

Nous réécrivons ces conditions et finalement constatons que

$$\begin{aligned}
&\{x_0^{**} = (x_{01}, x_{02}) \mid \left\{ \begin{array}{l} x_{02} d_2 \leq \sqrt{2}\, |d_2| \\ x_{01} d_1 + x_{02} d_2 \leq \sqrt{d_1^2 + d_2^2} \quad \forall\ d_1 \neq 0 \end{array} \right\} \ \forall\ d_2\} \\
&= \overline{B}
\end{aligned}$$

En effet, considérons un point $x_0^{**} = (x_{01}, x_{02})$ de $\overline{B}$ et montrons la première inégalité

a) Si $d_2 \leq 0$:

Utilisons le fait que

$$x_{02} \geq -1 \geq -\sqrt{2}$$

pour déduire

$$x_{02} d_2 \leq -\sqrt{2}\, d_2 = \sqrt{2}\, |d_2|$$

b) Si $d_2 > 0$:

Le résultat s'obtient tout aussi aisément en constatant que

$$x_{02} \leq 1 \leq \sqrt{2}$$

Concernant la deuxième inégalité, nous avons que

$$\begin{aligned} \forall\ d_1 > 0\ \ \forall\ d_2\ \ x_{01}d_1 + x_{02}d_2 &= \ < x_0^{**}, d > \\ &\leq \ ||x_0^{**}||\ ||d|| \end{aligned}$$

par l'inégalité de Cauchy-Schwarz.
Puisque $x_0^{**} \in \overline{B}$, nous obtenons alors

$$< x_0^{**}, d > \leq \ ||d|| \ = \sqrt{d_1^2 + d_2^2}$$

A présent, établissons la seconde inclusion.
Soit un point x_0^{**} vérifiant les deux inégalités ci-dessus. Envisageons l'alternative suivante :

a) Si $x_{01} \neq 0$:

Choisissons dans ce cas $d_1 = x_{01}$ et $d_2 = x_{02}$.
Nous avons dès lors,

$$x_{01}^2 + x_{02}^2 \leq \sqrt{x_{01}^2 + x_{02}^2}$$

c'est-à-dire

$$0 \leq \sqrt{x_{01}^2 + x_{02}^2} \leq 1$$

ou encore

$$x_{01}^2 + x_{02}^2 - 1 \leq 0$$

Ainsi, $x_0^{**} \in \overline{B}$.

b) Si $x_{01} = 0$:

Prenons cette fois $d_2 = x_{02}$. Par conséquent,

$$x_{02}^2 \leq \sqrt{d_1^2 + d_2^2} \quad \forall\ d_1 \neq 0$$

D'où, par passage à la limite lorsque d_1 tend vers 0,

$$x_{02}^2 \leq |x_{02}|$$

ce qui nous donne

$$|x_{02}| \leq 1$$

et donc, $x_0^{**} \in \overline{B}$. ∎

Nous observons donc bien que

$$\partial^2 f_1(x_0) + \partial^2 f_2(x_0) = [-1, 1] \times \{0\}$$

et

$$\sqrt{2}\, Co\{\partial^2 f_1(x_0) \cup \partial^2 f_2(x_0)\} = [-\sqrt{2}, \sqrt{2}] \times \{0\}$$

ne contiennent pas $\partial^2(f_1 + f_2)(x_0)$. ∎

Remarquons encore que, même dans le cas où f_1 et f_2 sont des fonctions deux fois différentiables, on n'a pas l'égalité

$$\partial^2(f_1 + f_2)(x_0) = \partial^2 f_1(x_0) + \partial^2 f_2(x_0)$$

De fait, le différentiel second $\partial^2(f_1 + f_2)(x_0)$ qui n'est autre que l'ellipsoïde $[\nabla^2(f_1+f_2)(x_0)]^{1/2}(\bar{B})$ associé à la matrice hessienne $\nabla^2(f_1+f_2)(x_0) = \nabla^2 f_1(x_0)+\nabla^2 f_2(x_0)$, est inclus (généralement strictement) dans l'addition

$$\partial^2 f_1(x_0) + \partial^2 f_2(x_0) = [\nabla^2 f_1(x_0)]^{1/2}(\bar{B}) + [\nabla^2 f_2(x_0)]^{1/2}(\bar{B})$$

Le caractère strict de cette inclusion provient du fait qu'en général, on a

$$\sqrt{a+b} < \sqrt{a} + \sqrt{b} \ \ \forall\ a \geq 0, b \geq 0$$

au lieu de l'égalité, celle-ci se produisant dans le cas particulier où $a = 0$ ou $b = 0$.
Ainsi, supposons que l'une des deux fonctions, soit f_2, soit polyédrale.
Nous obtenons dès lors que

$$\begin{aligned}\partial^2(f_1 + f_2)(x_0) &= \partial^2 f_1(x_0) + \partial^2 f_2(x_0)\\ &= \partial^2 f_1(x_0) + \{0\}\\ &= \partial^2 f_1(x_0)\end{aligned}$$

puisque

$$\begin{aligned}\sqrt{(f_1 + f_2)''(x_0, d)} &= \sqrt{f_1''(x_0; d) + f_2''(x_0; d)}\\ &= \sqrt{f_1''(x_0; d) + 0}\\ &= \sqrt{f_1''(x_0; d)}\\ &= \sqrt{f_1''(x_0; d)} + \sqrt{f_2''(x_0; d)}\end{aligned}$$

∎

3 Le maximum de fonctions

Considérons une famille finie $f_1, \cdots, f_m : I\!R^n \to I\!R$ de fonctions convexes pour lesquelles la dérivée directionnelle du second ordre, en $x_0 \in I\!R^n$ et dans la direction d, existe et est finie.
Soit alors

$$f(x) = \underset{1 \leq i \leq m}{\mathrm{Max}}\, f_i(x)$$

Nous savons que

$$f'(x_0; d) = \max_{i \in I(x_0)} f_i'(x_0; d) \text{ et } \partial f(x_0) = Co\{ \bigcup_{i \in I(x_0)} \partial f_i(x_0)\}$$

où $I(x_0) = \{i |\ f_i(x_0) = f(x_0)\}$.
Concernant la dérivée directionnelle du second ordre de f, il est aussi connu que

$$f''(x_0; d) = \max_{i \in I(x_0; d)} f_i''(x_0; d) \tag{24}$$

où $I(x_0; d) = \{i \in I(x_0) |\ f_i'(x_0; d) = f'(x_0; d)\}$.
Dès lors, f devient à son tour une fonction convexe [cfr A.I.10], définie sur $I\!R^n$ et telle que sa dérivée directionnelle du second ordre $f''(x_0; d)$ existe et est finie.

Nous donnons, à présent, une estimation inférieure et supérieure de l'ensemble $\partial^2 f(x_0)$ en termes des $\partial^2 f_i(x_0)$. Cette dernière estimation nécessite la convexité des fonctions $f_i''(x_0; .)$ lorsque $i \in I(x_0)$.

Proposition 3.1 :

Nous avons que

$$\bigcap_{i \in I(x_0)} \partial^2 f_i(x_0) \subseteq \partial^2 f(x_0)$$

De plus, si tous les $f_i''(x_0; .), i \in I(x_0)$ sont convexes, alors

$$\partial^2 f(x_0) \subseteq Co\{ \bigcup_{i \in I(x_0)} \partial^2 f_i(x_0)\} \tag{25}$$

Démonstration :

Au moyen de (24), nous établissons que

$$\begin{aligned}\sqrt{f''(x_0;d)} &= \sqrt{\max_{i\in I(x_0;d)} f_i''(x_0;d)} = \max_{i\in I(x_0;d)} \sqrt{f_i''(x_0;d)} \\ &\geq \min_{i\in I(x_0;d)} \sqrt{f_i''(x_0;d)}\end{aligned}$$

Comme $I(x_0;d) \subseteq I(x_0)$, nous avons encore

$$\begin{aligned}\sqrt{f''(x_0;d)} &\geq \min_{i\in I(x_0;d)} \sqrt{f_i''(x_0;d)} \\ &\geq \min_{i\in I(x_0)} \sqrt{f_i''(x_0;d)} \ \forall\, d\end{aligned}$$

Soit

$$x_0^{**} \in \bigcap_{i\in I(x_0)} \partial^2 f_i(x_0)$$

Nous obtenons alors

$$< x_0^{**}, d > \leq \sqrt{f_i''(x_0;d)} \ \forall\, d\, \forall\, i \in I(x_0)$$

c'est-à-dire

$$< x_0^{**}, d > \leq \min_{i\in I(x_0)} \sqrt{f_i''(x_0;d)} \ \forall\, d$$

D'où,

$$< x_0^{**}, d > \leq \sqrt{f''(x_0;d)} \ \forall\, d$$

ce qui signifie que $x_0^{**} \in \partial^2 f(x_0)$.
Reprenons à nouveau l'égalité (24) pour déduire que

$$\begin{aligned}\sqrt{f''(x_0;d)} &= \max_{i\in I(x_0;d)} \sqrt{f_i''(x_0;d)} \\ &\leq \max_{i\in I(x_0)} \sqrt{f_i''(x_0;d)} \ \forall\, d\end{aligned}$$

Considérons alors les sous-différentiels au point 0 de chacune des deux fonctions situées de part et d'autre de l'inégalité.

$$\begin{aligned}\partial\sqrt{f''(x;d)}(0) &= \{x^* | \sqrt{f''(x_0;d)} \geq < x^*, d > \ \forall\, d\} \\ &\subseteq \{x^* | < x^*, d > \leq \max_{i\in I(x_0)} \sqrt{f_i''(x_0;d)} \ \forall\, d\} \\ &= \partial\{\max_{i\in I(x_0)} \sqrt{f_i''(x_0;d)}\}(0) \\ &= Co\{\bigcup_{i\in I(x_0)} \partial\sqrt{f_i''(x_0;d)}(0)\} \\ &= Co\{\bigcup_{i\in I(x_0)} \partial^2 f_i(x_0)\}\end{aligned}$$

D'où l'inclusion annoncée. ■

L'exemple que nous considérons maintenant montre que sans l'hypothèse de convexité des $f_i''(x_0;.)$, l'inclusion (25) peut ne pas se vérifier.

Exemple 6 : Soit la fonction $f = \max\{f_1, f_2\}$ avec

$$\begin{aligned} f_1(x) &= \text{Max}\{\frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + x_1, 0\} \\ f_2(x) &= \text{Max}\{\frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + x_2, 0\} \end{aligned}$$

Nous avons déjà calculé, dans l'exemple 5,

$$\partial^2 f_1(x_0) = [0,1] \times \{0\} \quad \text{et} \quad f_1'(x_0; d) = \max(0, d_1)$$

où $x_0 = (0,0)$.
Nous obtenons donc de façon similaire (observer le rôle joué par x_1 et x_2 dans chacune des fonctions $f_i(x)$ $i = 1,2$)

$$\partial^2 f_2(x_0) = \{0\} \times [0,1] \quad \text{et} \quad f_2'(x_0; d) = \max(0, d_2)$$

Dès lors,

$$Co[\partial^2 f_1(x_0) \cup \partial^2 f_2(x_0)] = \{(z_1, z_2) \in I\!R_+^2 \mid z_1 + z_2 \leq 1\}$$

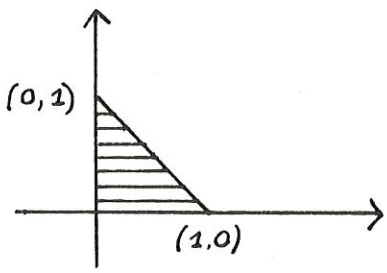


Figure 23

D'autre part, recherchons $\partial^2 f(x_0)$.
Nous constatons que

$$\begin{aligned} I(x_0) &= \{i \mid f_i(x_0) = f(x_0) = 0\} \\ &= \{1,2\} \end{aligned}$$

ce qui nous donne

$$\begin{aligned} f'(x_0;d) &= \max_{i=1,2} f_i'(x_0;d) \\ &= \begin{cases} 0 & \text{si } d_1 \leq 0 \quad \text{et } d_2 \leq 0 \\ d_2 & \text{si } d_1 \leq 0 \quad \text{et } d_2 > 0 \\ d_1 & \text{si } d_1 > 0 \quad \text{et } d_2 \leq 0 \\ \max(d_1, d_2) & \text{si } d_1 > 0 \quad \text{et } d_2 > 0 \end{cases} \end{aligned}$$

Dès lors,

$$f''(x_0;d) = \max_{i \in I(x_0;d)} f_i''(x_0;d)$$

c'est-à-dire,

$$\begin{cases} \text{si } d_1 \leq 0 \text{ et } d_2 \leq 0, f''(x_0;d) \text{ vaut } \max(f_1''(x_0;d), f_2''(x_0;d)) \\ \text{si } d_1 \leq 0 \text{ et } d_2 > 0, f''(x_0;d) \text{ vaut } f_2''(x_0;d) \\ \text{si } d_1 > 0 \text{ et } d_2 \leq 0, f''(x_0;d) \text{ vaut } f_1''(x_0;d) \\ \text{si } d_1 > 0 \text{ et } d_2 > 0, \text{ alors } f''(x_0;d) \text{ vaut} \end{cases}$$

$$\begin{cases} f_1''(x_0;d) & \text{dans le cas où } d_1 > d_2 \\ f_2''(x_0;d) & \text{dans le cas où } d_2 > d_1 \\ \text{Max}(f_1''(x_0;d), f_2''(x_0;d)) & \text{dans le cas où } d_1 = d_2 \end{cases}$$

et, en reprenant les valeurs trouvées dans l'exemple 5 pour $f_1''(x_0;d)$, à savoir :

$$f_1''(x_0;d) = \begin{cases} d_1^2 + d_2^2 & \text{si } d_1 \geq 0 \\ 0 & \text{si } d_1 < 0 \end{cases}$$

nous déduisons

$$f_2''(x_0;d) = \begin{cases} d_1^2 + d_2^2 & \text{si } d_2 \geq 0 \\ 0 & \text{si } d_2 < 0 \end{cases}$$

et

$$f''(x_0;d) = \begin{cases} 0 & \text{si } d_1 < 0 \text{ et } d_2 < 0 \\ d_1^2 & \text{si } d_1 < 0 \text{ et } d_2 = 0 \\ d_2^2 & \text{si } d_1 = 0 \text{ et } d_2 < 0 \\ 0 & \text{si } d_1 = 0 \text{ et } d_2 = 0 \\ d_1^2 + d_2^2 & \text{sinon} \end{cases}$$

Finalement, nous obtenons

$$\partial^2 f(x_0) = \{z| < z, d > \leq \begin{cases} 0 & \text{si } d_1 < 0 \text{ et } d_2 < 0 \text{ ou} \\ & \text{si } d_1 = d_2 = 0 \\ -d_1 & \text{si } d_1 < 0 \text{ et } d_2 = 0 \; \} \\ -d_2 & \text{si } d_2 < 0 \text{ et } d_2 = 0 \\ \sqrt{d_1^2 + d_2^2} & \text{sinon} \end{cases}$$

La condition (26) implique alors $z_1 \geq 0$ et $z_2 \geq 0$ avec $z = (z_1, z_2)$.
Les conditions (27), (28) et (29) sont redondantes.
La condition (30) permet d'établir que $(z_1, z_2) \in \bar{B}$ (preuve tout à fait similaire à celle de l'exemple 5).
Dès lors,

$$\partial^2 f(x_0) = \{(z_1, z_2) \in I\!R_+^2 | \; z_1^2 + z_2^2 \leq 1\}$$

Nous constatons ainsi que

$$Co[\partial^2 f_1(x_0) \cup \partial^2 f_2(x_0)] \subset \partial^2 f(x_0)$$

où l'inclusion est stricte. ■

De la proposition 3.1, nous déduisons immédiatement la propriété suivante.

<u>**Corollaire 3.2 :**</u>

Supposons que toutes les fonctions f_i pour $i \in I(x_0)$ soient deux fois différentiables au point x_0 et notons E_i les ellipsoïdes associés aux $\nabla^2 f_i(x_0)$. Alors, nous avons

$$\bigcap_{i \in I(x_0)} E_i \subseteq \partial^2 f(x_0) \subseteq Co\{\bigcup_{i \in I(x_0)} E_i\}$$

Comme illustration de ce résultat, reprenons l'exemple 3 dans lequel f représente le maximum de deux fonctions quadratiques convexes sur $I\!R^2$.
Comme de telles fonctions sont deux fois différentiables, nous pouvons identifier

$$\begin{aligned} \partial^2 f_1(x_0) &= [\nabla^2 f_1(x_0)]^{1/2}(\bar{B}) \\ \text{et} \quad \partial^2 f_2(x_0) &= [\nabla^2 f_2(x_0)]^{1/2}(\bar{B}) \end{aligned}$$

Nous observons ainsi, au moyen de la figure 19, que $\partial^2 f(x_0)$ contient $\partial^2 f_1(x_0) \cap \partial^2 f_2(x_0)$ et est contenu dans l'enveloppe convexe de $\partial^2 f_1(x_0) \cup \partial^2 f_2(x_0)$.

CHAPITRE IV

DERIVEES SECONDES INTERIEURE

ET EXTERIEURE

Commençons par rappeler quelques résultats fondamentaux concernant la géométrie des ellipsoïdes dans $I\!R^n$.

1 Propriétés des ellipsoïdes

La notion géométrique associée à une matrice M symétrique, semi-définie positive, de dimension $n \times n$, est l'ellipsoïde $E = M^{1/2}(\bar{B})$ où $\bar{B}$ désigne à nouveau la boule unité fermée de $I\!R^n$. La fonction d'appui de l'ellipsoïde est alors obtenue par

$$\begin{aligned} \delta^*(E|d) &= \sup_{c \in E} < c, d > \\ &= \sup_{||b|| \leq 1} < M^{1/2} b, d > \\ &= \sup_{||b|| \leq 1} < b, M^{1/2} d > \\ &= \frac{< M^{1/2} d, M^{1/2} d >}{||M^{1/2} d||} \\ &= ||M^{1/2} d|| = \sqrt{< M^{1/2} d, M^{1/2} d >} \\ &= \sqrt{< M d, d >} \ \ \forall\, d \in I\!R^n \end{aligned} \tag{31}$$

Réciproquement, un ellipsoïde E de centre $0 \in I\!R^n$ peut se représenter par $E = N(\bar{B})$ pour une certaine matrice N carrée de dimension n. Si N_1 et N_2 sont deux matrices associées du même ellipsoïde, c'est-à-dire $E = N_1(\bar{B}) = N_2(\bar{B})$, nous obtenons

$$N_1 N_1^T = N_2 N_2^T$$

En effet, commençons par démontrer le résultat préliminaire suivant.

$$\forall\, x \in I\!R^n \quad x^T N_1 N_1^T x = x^T N_2 N_2^T x \qquad (*)$$

Considérons ainsi un point x de $I\!R^n$. Sans perte de généralité, nous pouvons supposer qu'il est non nul. Nous avons

$$\delta^*(N_1(\overline{B})|\, x) = \sup_{b \in \overline{B}} x^T N_1 b = x^T N_1 \frac{N_1^T x}{||N_1^T x||} = ||N_1^T x||$$

et de façon identique

$$\delta^*(N_2(\overline{B})|\, x) = ||N_2^T x||$$

Dès lors, puisque $N_1(\overline{B}) = N_2(\overline{B})$, nous obtenons

$$||N_1^T x|| = ||N_2^T x||$$

et donc

$$||N_1^T x||^2 = ||N_2^T x||^2$$

c'est-à-dire $\quad x^T N_1 N_1^T x = x^T N_2 N_2^T x$.
De plus, notant les matrices symétriques

$$\begin{aligned} N_1 N_1^T &= [a_{ij}]_{i,j} \\ N_2 N_2^T &= [b_{ij}]_{i,j} \end{aligned}$$

nous montrons

a) $a_{ii} = b_{ii} \quad \forall\, i$:
De fait, nous appliquons $(*)$ avec $x = e_i$ où e_i représente le $i^{\text{ème}}$ vecteur de la base canonique de $I\!R^n$, c'est-à-dire

$$e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$$

avec le 1 situé à la $\mathrm{i}^{\text{ème}}$ place.
Nous avons ainsi

$$x^T N_1 N_1^T x = (0, \cdots, 1, \cdots, 0)[a_{ij}] \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

$$= \quad (a_{i1}, \cdots, a_{in}) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a_{ii}$$

et de ma même manière

$$x^T N_2 N_2^T x = b_{ii}$$

Ce résultat est valable pour tout i , $1 \leq i \leq n$.

b) $a_{ij} = b_{ij} \quad \forall\, i \neq j$.

Cette fois, nous appliquons $(*)$ avec

$$x = (0, \cdots, 1, 0, \cdots, 0, 1, 0, \cdots, 0)$$

où les 1 situés à la $i^{\text{ème}}$ et la $j^{\text{ème}}$ place.

Nous obtenons

$$\begin{aligned} x^T N_1 N_1^T x &= (a_{i1} + a_{j1}, \cdots, a_{in} + a_{jn}) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \\ &= a_{ii} + a_{ji} + a_{ij} + a_{jj} \\ &= a_{ii} + a_{jj} + 2a_{ij} \end{aligned}$$

et $x^T N_1 N_2^T x = b_{ii} + b_{jj} + 2b_{ij}$

en vertu du caractère symétrique de $N_1 N_1^T$ et $N_2 N_2^T$.

Consécutivement,

$$a_{ij} = b_{ij}$$

puisque $a_{ii} = b_{ii}$ et $a_{jj} = b_{jj}$ [voir a)] ■

Ainsi,

$$E = M^{1/2}(\bar{B}) \text{ où } M = M^{1/2}M^{1/2} = NN^T$$

est déterminée de manière unique.
Dès lors, nous pouvons établir une bijection entre l'ensemble $\mathcal{E}$ des ellipsoïdes de $\mathbb{R}^n$, centrés en l'origine, et l'ensemble $\mathcal{P}_n$ des matrices symétriques, semi-définies positives, de dimension $n \times n$.

$$\begin{array}{rcl} \mathcal{E} & & \mathcal{P}_n \\ E = N(\bar{B}) & \longrightarrow & M = NN^T \\ E = M^{1/2}(\bar{B}) & \longleftarrow & M \end{array}$$

Lorsque M est définie positive, donc régulière, nous avons déjà noté [cfr proposition I.2.4.2.1] que

$$E = \{x \mid < M^{-1}x, x > \leq 1\}$$

Nous citons à présent un résultat établissant la connexion entre les propriétés géométriques de l'ellipsoïde E et les caractéristiques algébriques de la matrice M.

Proposition 1.1 :

Le diamètre de E, c'est-à-dire la longueur du plus grand de ses axes, est égal à $2\sqrt{\Lambda}$ où Λ désigne la plus grande valeur propre de M (positive puisque M est semi-définie positive). La largeur de E, c'est-à-dire la longueur du plus petit axe, vaut $2\sqrt{\lambda}$ où λ est la plus petite valeur propre de M.
De plus, les axes de symétrie de E correspondent aux vecteurs propres de M.

Consécutivement, si M est singulière, E n'est autre que l'illipsoïde "plat".

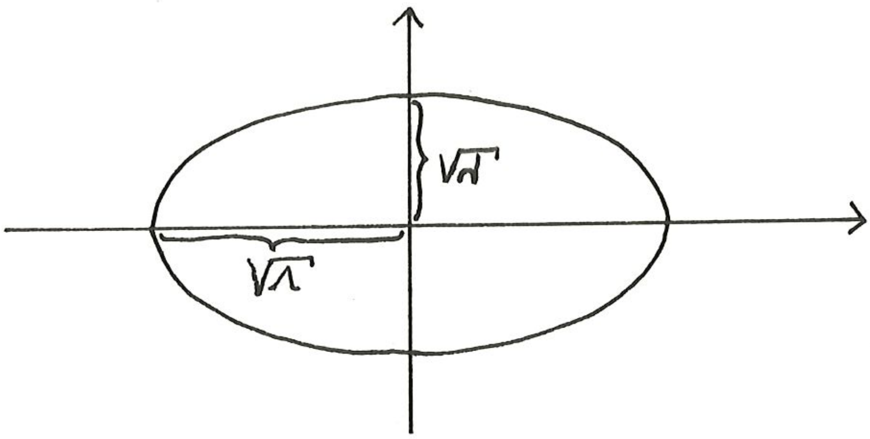


Figure 24

Nous énonçons une dernière propriété concernant le volume de l'ellipsoïde E, noté $vol\,(E)$.

Proposition 1.2 :

$$vol\,(E) = \sqrt{det\ M} \cdot vol\,(\bar{B})$$

2 Dérivées secondes intérieure et extérieure ou comment approcher $\partial^2 f(x_0)$ au moyen d'ellipsoïdes

Un résultat dû à Alexandroff (1939) affirme qu'une fonction f définie sur $\mathbb{R}^n$ est deux fois différentiable presque partout. Ainsi, pour presque tout point x_0 de $\mathbb{R}^n$, f admet un développement de Taylor du second ordre en ce sens qu'il existe une matrice symétrique, semi-définie positive $\nabla^2 f(x_0)$ vérifiant

$$\begin{aligned} f(x) &= f(x_0) + < \nabla f(x_0), x - x_0 > + \frac{1}{2} < \nabla^2 f(x_0)(x - x_0), x - x_0 > \\ &\quad + o(||x - x_0||^2) \end{aligned}$$

Pour un tel point x_0, nous avons $f''(x_0; d) = < \nabla^2 f(x_0)d, d >$ [cfr paragraphe I.2.1] et

$$\partial^2 f(x_0) = E = [\nabla^2 f(x_0)]^{1/2}(\bar{B})$$

comme vu précédemment dans la proposition II.2.5.

Par contre, lorsque l'on considère un point x_0 pour lequel $\partial^2 f(x_0)$ n'est plus un ellipsoïde, il n'y a a priori, aucune raison d'associer à cet ensemble une unique matrice, symétrique et semi-définie positive. Cependant, nous pourrons approcher le différentiel second au moyen de deux éléments choisis respectivement dans l'ensemble $\overline{\mathcal{E}}$ des ellipsoïdes centrés en 0 contenant $\partial^2 f(x_0)$ et l'ensemble $\underline{\mathcal{E}}$ des ellipsoïdes de centre 0 contenu dans $\partial^2 f(x_0)$. Le critère choisi pour prendre l'élément maximal dans $\underline{\mathcal{E}}$ et l'élément minimal dans $\overline{\mathcal{E}}$ est le volume. Nous prouvons à présent l'existence et l'unicité de tels ellipsoïdes.

Proposition 2.1 :

Il existe un et un seul ellipsoïde de volume maximal dans $\underline{\mathcal{E}}$, soit $\underline{E}$ et il existe un et un seul ellipsoïde de volume minimal dans $\overline{\mathcal{E}}$, noté $\overline{E}$.

Démonstration :

Pour simplifier la preuve, nous supposons que $\partial^2 f(x_0)$ est de dimension pleine (Si ce n'est pas le cas, le même raisonnement pourra s'effectuer dans l'enveloppe affine de $\partial^2 f(x_0)$).
Utilisant la proposition 1.2, nous observons que le volume de l'ellipsoïde $E = M^{1/2}(\overline{B})$ est proportionnel à $(det\ M)^{1/2}$. Ainsi, en supposant les matrices associées aux ellipsoïdes $\underline{E}$ et $\overline{E}$ définies positives (cas de dégénérescence écarté), nous sommes mis en présence des deux problèmes d'optimisation suivants :

$$(\underline{P}) \equiv \begin{cases} \text{Maximiser } (det\ M)^{1/2} \\ \text{sous contraintes } M \in \dot{\mathcal{P}}_n \text{ et } \{x \mid\ < M^{-1}x, x > \leq 1\} \subseteq \partial^2 f(x_0) \end{cases}$$

$$(\overline{P}) \equiv \begin{cases} \text{Minimiser } (det\ M)^{1/2} \\ \text{sous contraintes } M \in \dot{\mathcal{P}}_n \text{ et } \partial^2 f(x_0) \subseteq \{x \mid\ < M^{-1}x, x > \ |eq1\} \end{cases}$$

où $\dot{\mathcal{P}}_n$ dsigne ici l'ensemble des matrices symétriques, définies positives, de dimension $n \times n$.
Penchons-nous tout d'abòrd sur le problème $(\underline{P})$ et montrons la convexité de l'ensemble $\dot{\mathcal{P}}_n$.
En effet, considérons $\alpha A + (1-\alpha)B$ où $A, B \in \dot{\mathcal{P}}_n$ et $\alpha \in [0,1]$. Nous obtenons

a) le caractère symétrique de cette expression car

$$\begin{aligned}(\alpha A)^T &= \alpha A^T = \alpha A \\ (A+B)^T &= A^T + B^T = A + B\end{aligned}$$

b) le caractère défini positif puisque

$$\begin{aligned} < [\alpha A + (1-\alpha)B]d, d > &= \alpha < Ad, d > + (1-\alpha) < Bd, d > \\ &> 0\end{aligned}$$

De plus, si

$$\{x \mid \, < M^{-1}x, x > \leq 1\} = M^{1/2}(\bar{B}) \subseteq \partial^2 f(x_0)$$

et

$$\{x \mid \, < N^{-1}x, x > \leq 1\} = N^{1/2}(\bar{B}) \subseteq \partial^2 f(x_0)$$

Alors

$$\begin{aligned}\{x \mid \, < [\alpha M + (1-\alpha)N]^{-1}x, x > \leq 1\} &= [\alpha M + (1-\alpha)N]^{1/2}(\bar{B}) \\ &\subseteq \partial^2 f(x_0)\end{aligned}$$

pour tout $\alpha \in [0, 1]$.
De fait, nous pouvons écrire

$$\begin{aligned} E &= [\alpha M + (1-\alpha)N]^{1/2}(\bar{B}) \\ &= \{x \mid \, < x, d > \leq \delta^*(E|d) \;\; \forall \, d\} \end{aligned} \tag{32}$$

Prenons alors un point arbitraire e de E. Nous avons

$$\begin{aligned} < e, d > &\leq \delta^*(E|d) = \sup_{x \in E} \; < x, d > \\ &= \sqrt{d^T[\alpha M + (1-\alpha)N]d} \qquad \text{[cfr (31)]} \\ &= \sqrt{\alpha d^T M d + (1-\alpha) d^T N d}\end{aligned}$$

Or,

$$\begin{aligned} M^{1/2}(\bar{B}) &= \{x \mid \, < x, d > \leq \delta^*(M^{1/2}(\bar{B})|d) \;\; \forall \, d\} \\ &= \{x \mid \, < x, d > \leq \sqrt{d^T M d} \;\; \forall \, d\} \qquad \text{[cfr (31)]} \\ &\subseteq \partial^2 f(x_0) = \{x \mid \, < x, d > \leq \sqrt{f''(x_0; d)} \;\; \forall \, d\}\end{aligned}$$

en vertu de la proposition II.2.1.
Donc,

$$\sqrt{d^T M d} \leq \sqrt{f''(x_0; d)} \ \ \forall \ d$$

et, de façon identique

$$\sqrt{d^T N d} \leq \sqrt{f''(x_0; d)} \ \ \forall \ d$$

Dès lors,

$$\begin{aligned} < e, d > &= \sqrt{\alpha d^T M d + (1-\alpha) d^T N d} \\ &\leq \sqrt{\alpha f''(x_0; d) + (1-\alpha) f''(x_0; d)} \\ &= \sqrt{f''(x_0; d)} \quad \forall \ d \end{aligned}$$

et ainsi, $e \in \partial^2 f(x_0)$, d'où l'inclusion annoncée.
Il suit que l'ensemble des contraintes (sur M) dans $(\underline{P})$ est convexe.
Par ailleurs, maximiser $(det\ M)^{1/2}$ dans $(\underline{P})$ revient à maximiser $\log(det\ M)$.
Nous montrons la concavité stricte de la fonction $M \to \log(det M)$ afin d'en déduire l'unicité de $\underline{E}$.
Nous faisons, dans ce but, appel au théorème 6 de [6, p. 63]. Celui-ci nous permet d'obtenir

$$det\ [\lambda A + (1-\lambda)B] \geq (det\ A)\lambda (det\ B)^{1-\lambda} \ \ \forall \ \lambda \in [0,1]$$

puisque A et B constituent des matrices symétriques définies positives.
Ainsi,

$$log\ det\ [\lambda A + (1-\lambda)B] \geq \lambda log\ det\ A + (1-\lambda) log\ det\ A \ \ \forall \ \lambda \in [0,1]$$

c'est-à-dire

$$log\ det\ [\lambda A + (1-\lambda)B] > \lambda log\ det\ (A) + (1-\lambda) log\ det\ (B) \ \ \forall \ \lambda \in]0,1[$$

lorsque A et B sont choisies distinctes.
Pour prouver l'existence de $\underline{E}$, nous considérons une matrice $\overline{M} \in \dot{\mathcal{P}}_n$ telle que

$$\{x| \ < (\overline{M})^{-1} x, x > \leq 1\} \subseteq \partial^2 f(x_0)$$

Une telle matrice existe car, $\partial^2 f(x_0)$ étant de dimension pleine et contenant l'origine [cfr proposition II.2.2],

$$\exists r > 0 \ \ \overline{B}(0,r) \subseteq \partial^2 f(x_0)$$

Or,

$$\overline{B}(0,r) = \{x| \ < (r^2I)^{-1}x, x > \leq 1\}$$

où I désigne la matrice unité.
Nous pouvons dès lors prendre $\overline{M} = r^2 I$, qui constitue bien une matrice symétrique, définie positive. En conséquence, nous avons

$$\alpha \overset{\text{not}}{=} det\ \overline{M} > 0$$

Ainsi, notre problème $(\underline{P})$ est équivalent à

$$(\underline{P})^* \equiv \begin{cases} \text{Maximiser } (det\ M)^{1/2} & \\ \text{sous contraintes} & M \in \dot{\mathcal{P}}_n \\ & \{x| \ < M^{-1}x, x > \leq 1\} \subseteq \partial^2 f(x_0) \\ & det\ M \geq \alpha \end{cases}$$

Montrons que $(\underline{P})^*$ admet au moins une solution.
Nous savons tout d'abord que

$$M \to (det\ M)^{1/2}$$

est continue.
D'autre part, nous prouvons que

$$\mathcal{A} = \{M| \ M \text{ admissible pour } (\underline{P})^*\}$$

est compact.

a) $\mathcal{A}$ est fermé :
en effet, considérons la suite de matrices $(M_n) \subseteq \dot{\mathcal{P}}_n$ convergente vers M et telle que

$$[M_n]^{1/2}(\overline{B}) \subseteq \partial^2 f(x_0) \ \ \forall\ n$$

et

$$det\ M^n \geq \alpha$$

Nous désirons montrer que M est admissible pour $(\underline{P})^*$.

(i) M est symétrique car

$$M^T = (\lim_{n\to\infty} M^n)^T = \lim_{n\to\infty} (M)^T = \lim_{n\to\infty} M_n = M$$

(ii) M est semi-définie positive puisque

$$< Md, d >= \lim_{n\to\infty} < M_n d, d > \geq 0$$

par continuité du produit scalaire.

(iii) Nous avons que

$$M^{1/2}(\overline{B}) \subseteq \partial^2 f(x_0)$$

De fait, soit $z \in M^{1/2}(\overline{B})$. Nous pouvons ainsi écrire, pour $b \in \overline{B}$,

$$\begin{aligned} z &= M^{1/2}(b) = (\lim_{n\to\infty} M_n)^{1/2} b \\ &= \lim_{n\to\infty} (M_n)^{1/2} b \end{aligned}$$

Posons $z_n = [M_n]^{1/2} b \in [M_n]^{1/2}(\overline{B})$. Nous obtenons dès lors que

$$\lim_{n\to\infty} z_n = z$$

et puisque $\partial^2 f(x_0)$ est fermé [cfr proposition II.2.2], z constitue bien un de ses éléments.

(iv) Comme $M_n \to det\ M_n$ est continue, nous avons

$$det\ M = \lim_{n\to\infty} det\ M_n \geq \alpha > 0$$

D'où, M est aussi définie positive.

b) $\mathcal{A}$ est borné :

Nous cherchons donc à établir l'existence d'une cons.ante $c > 0$ tel que
$\forall\ M \in \mathcal{A} \quad ||M|| \leq c$.
Nous savons que

$$\begin{aligned} ||M|| &= \sup_{||x||\leq 1} < Mx, x > \\ &= \sup_{||x||\leq 1} < M^{1/2}x, M^{1/2}x > \\ &= \sup_{||x||\leq 1} ||M^{1/2}x||^2 \\ &\leq k^2 = c \end{aligned}$$

puisque $M^{1/2}(\bar{B}) \leq \partial^2 f(x_0)$ et que ce dernier ensemble est borné.

Considérons à présent le problème $(\overline{P})$ et prouvons l'existence de $\overline{E}$.
Pour ce faire, nous prenons une matrice $\overline{M} \in \dot{\mathcal{P}}_n$ telle que

$$\partial^2 f(x_0) \subseteq \{x| \ < (\overline{M})^{-1}x, x > \leq 1\}$$

Une telle matrice existe car, en vertu du caractère borné de $\partial^2 f(x_0)$,

$$\exists R > 0 \text{ tel que}$$

$$\partial^2 f(x_0) \subseteq \overline{B}(0, R) = \{x| \ < \frac{1}{R^2}Ix, x > \leq 1\} = (R^2 I)^{1/2}(\overline{B})$$

Nous pouvons alors poser $\overline{M} = R^2 I$. Il s'agit bien d'une matrice symétrique, définie positive. Il s'en suit que

$$\alpha \overset{\text{not}}{=} det\ \overline{M} > 0$$

D'autre part, puisque $\partial^2 f(x_0)$ est supposée de dimension pleine,
$\exists r > 0$ tel que

$$\overline{B}(0, r) \subseteq \partial^2 f(x_0)$$

Dès lors, pour toute matrice M admissible pour $(\overline{P})$, nous avons

$$vol\ [\overline{B}(0, r)] \leq vol\ \{x| \ < M^{-1}x, x > \leq 1\}$$

c'est-à-dire, en vertu de la proposition 1.2,

$$det\ r^2 I \leq det\ M$$

où

$$det\ r^2 I = (r^2)^n = \beta > 0$$

Consécutivement, le problème $(\overline{P})$ est équivalent à

$$(\overline{P})^* \equiv \begin{cases} \text{Minimiser } (det\ M)^{1/2} & \\ \text{sous contraintes} & M \in \dot{\mathcal{P}}_n \\ & \partial^2 f(x_0) \subseteq \{x| \ < M^{-1}x, x > \leq 1\} \\ & 0 < \beta \ \leq \ det\ M \leq \alpha \end{cases}$$

Montrons que $(\overline{P})^*$ admet au moins une solution.
Nous savons premièrement que

$$M \rightarrow (det\ M)^{1/2}$$

est continue.

Par ailleurs, nous prouvons que

$$\mathcal{B} = \{M |\ M \text{ admissible pour } (\overline{P})^*\}$$

est compact.

a) $\mathcal{B}$ est fermé :

De fait, considérons la suite de matrices $(M_n) \subseteq \dot{\mathcal{P}}_n$, convergente vers M et telle que

$$[M_n]^{1/2}(\overline{B}) \supseteq \partial^2 f(x_0)$$

et

$$det\ M_n \geq \beta$$

Nous souhaitons prouver que M est admissible pour $(\overline{P})^*$.

(i) M est symétrique car

$$M^T = (\lim_{n\to\infty} M_n)^T = \lim_{n\to\infty} (M_n)^T = \lim_{n\to\infty} M_n = M$$

(ii) M est semi-définie positive puisque

$$< Md, d > = \lim_{n\to\infty} \ < M_n d, d > \geq 0$$

par continuité du produit scalaire.

(iii) Nous avons que

$$\partial^2 f(x_0) \subseteq [M]^{1/2}(\overline{B})$$

En effet, soit $x^{**} \in \partial^2 f(x_0)$. Nous savons que

$$< (M_n)^{-1} x^{**}, x^{**} > \leq 1 \quad \forall\, n$$

et donc

$$< M^{-1} x^{**}, x^{**} > \leq 1$$

(iv) $M_n \to det\ M_n$ étant une fonction continue, nous avons

$$0 < \beta \leq \lim_{n\to\infty} \ det\ M_n = det\ M \leq \alpha$$

Ainsi, M est aussi définie positive.

b) $\mathcal{B}$ est borné :

Nous avons déjà établi que

$$\overline{B}(0,r) \subseteq \{x| < M^{-1}x, x > \leq 1\} \quad \forall\, M \in \mathcal{B}$$

et comme la racine carrée de la plus petite valeur propre de M représente la demi-longueur du plus petit axe de l'ellipsoïde [voir proposition 1.1], nous obtenons
$\forall\, M \in \mathcal{B}$ la plus petite valeur propre de $M \geq r^2 > 0$
Envisageons le raisonnement par l'absurde suivant.
Supposons que $\mathcal{B}$ ne soit pas borné.
Il existe dès lors une suite (M_k) dans $\mathcal{B}$ telle que

$$||M_k||_2 \to \infty$$

c'est-à-dire

$$\text{la plus grande valeur propre } \Lambda_k \text{ de } M_k \to +\infty$$

Par conséquent, puisque

$$\begin{aligned} det\ M_k &= \text{le produit de ses valeurs propres} \\ &\geq (n-1)r^2\Lambda_k \end{aligned}$$

nous obtenons que

$$det\ M_k \to +\infty$$

ce qui n'est pas possible car $det\ M_k \leq \alpha$.

Démontrons maintenant l'unicité de $\overline{E}$ en faisant de nouveau appel à un raisonnement par l'absurde.
Supposons donc que M et N sont deux solutions distinctes de $(\overline{P})$ et formons la combinaison suivante

$$[\alpha M^{-1} + (1-\alpha)N^{-1}]^{-1} \quad \forall\, \alpha \in]0,1[$$

Cette expression satisfait les contraintes imposées aux matrices de $(\overline{P})$. En effet,

(i) cette matrice est symétrique :

$$\begin{aligned} ([\alpha M^{-1} + (1-\alpha)N^{-1}]^{-1})^T &= ([\alpha M^{-1} + (1-\alpha)N^{-1}]^T)^{-1} \\ &= [\alpha(M^T)^{-1} + (1-\alpha)(N^T)^{-1}]^{-1} \\ &= [\alpha M^{-1} + (1-\alpha)N^{-1}]^{-1} \end{aligned}$$

puisque les matrices M et N sont elles-mêmes symétriques.

(ii) cette matrice est définie positive :

car si une matrice est définie positive, son inverse l'est encore et

$$\begin{aligned} < [\alpha M^{-1} + (1-\alpha)N^{-1}]d, d > &= \alpha < M^{-1}d, d > +(1-\alpha) < N^{-1}d, d > \\ &> 0 \end{aligned}$$

(iii) $\{x \mid < [\alpha M^{-1} + (1-\alpha)N^{-1}]x, x > \leq 1\} \supseteq \partial^2 f(x_0)$:

Soit un point x de $\partial^2 f(x_0)$. Comme M et N sont des éléments du domaine de $(\overline{P})$, nous avons

$$\{x \mid < M^{-1}x, x > \leq 1\} \supseteq \partial^2 f(x_0)$$

et

$$\{x \mid < N^{-1}x, x > \leq 1\} \supseteq \partial^2 f(x_0)$$

de sorte que

$$\begin{aligned} < [\alpha M^{-1} &+ (1-\alpha)N^{-1}]x, x > \\ &= \alpha < M^{-1}x, x > +(1-\alpha) < N^{-1}x, x > \\ &\leq \alpha \cdot 1 + (1-\alpha) \cdot 1 \\ &\leq 1 \end{aligned}$$

Par conséquent, l'inclusion est bien satisfaite.

De plus,

$log\ \{det\ [\alpha M^{-1} + (1-\alpha)N^{-1}]^{-1}\}$

$$\begin{aligned} &= log\ \{\frac{1}{det\ [\alpha M^{-1} + (1-\alpha)N^{-1}]}\} \\ &= -log\ det\ [\alpha M^{-1} + (1-\alpha)N^{-1}] \end{aligned}$$

et en utilisant le caractère strictement concave de la fonction $log\ det\ (\cdot)$, nous obtenons $-log\ det[\alpha M^{-1} + (1-\alpha)N^{-1}]$

$$\begin{aligned} &< -\alpha log\ det\ (M^{-1}) - (1-\alpha) log\ det\ (N^{-1}) \\ &= -\alpha log\ \frac{1}{det\ M} - (1-\alpha) log\ \frac{1}{det\ N} \\ &= \alpha log\ (det\ M) + (1-\alpha) log\ (det\ N) \end{aligned}$$

Finalement,

$$\begin{aligned} log\ \{det\ [\alpha M^{-1} + (1-\alpha)N^{-1}]^{-1}\} &< \alpha log\ (det\ M) + (1-\alpha) log\ (det\ N) \\ &= log\ (det\ M) \end{aligned}$$

puisque M et N sont solutions de $(\overline{P})$ et que trouver $\overline{E}$ revient à minimiser la fonction $log\,(det\ M)$. D'où la contradiction attendue. ∎

Les résultats précédents conduisent aux définitions suivantes.

Nous appellerons *dérivée seconde intérieure* (respectivement *dérivée seconde extérieure*) de f au point x_0, l'unique matrice symétrique, semi-définie positive, notée $\underline{\nabla}^2 f(x_0)$ (respectivement $\overline{\nabla}^2 f(x_0)$) associée à l'unique ellipsoïde de volume maximal (minimal), centré en l'origine et contenu dans (contenant) $\partial^2 f(x_0)$.
Ainsi,

$$\underline{E} = [\underline{\nabla}^2 f(x_0)]^{1/2}(\bar{B})$$

et

$$\overline{E} = [\overline{\nabla}^2 f(x_0)]^{1/2}(\bar{B})$$

Aux points où f est deux fois différentiable, nous obtenons évidemment que

$$\overline{\nabla}^2 f(x_0) = \underline{\nabla}^2 f(x_0) = \nabla^2 f(x_0)$$

Illustrons ces notions au moyen de l'exemple 1, étudié précédemment. Nous y prenions en considération la fonction

$$f(x_1, x_2) = \max(0, x_1^2 + x_2^2 + 2x_1)$$

et avions établi que

$$\partial^2 f(x_0) = [0, \sqrt{2}] \times \{0\}$$

pour $x_0 = (0, 0)$.

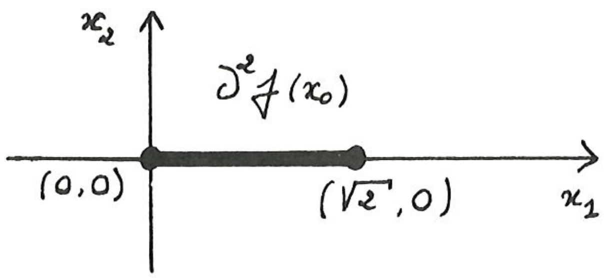


Figure 25

Calculons les matrices associées à $\underline{E}$ et $\overline{E}$.
Nous constatons que l'ellipsoïde de volume maximal, centré en l'origine et inclus dans

$\partial^2 f(x_0)$ n'est autre que l'origine. Dès lors,

$$\underline{\nabla}^2 f(x_0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

D'autre part, l'ellipsoïde de volume minimal, centré en 0 et contenant $\partial^2 f(x_0)$ est l'ellipsoïde plat dont le diamètre vaut $2\sqrt{2}$. Ainsi, ses axes de symétrie sont les vecteurs

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Comme les longueurs des axes fournissent (au facteur multiplicatif $\frac{1}{2}$ près) les racines carrées des valeurs propres [cfr proposition 1.1 dans le cas particulier où $I\!R^n = I\!R^2$], nous obtenons

$$\lambda_1 = \text{la plus petite valeur propre de } \overline{\nabla}^2 f(x_0) = 0$$
$$\lambda_2 = \text{la plus grande valeur propre de } \overline{\nabla}^2 f(x_0) = 2$$

D'où,

$$\overline{\nabla}^2 f(x_0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

∎

Nous clôturons ce chapitre en donnant une approximation quadratique de $f(x_0 + \lambda d)$ utilisant la dérivée seconde intérieure.

Proposition 2.2 :

Soit donné x_0^* dans $\partial f(x_0)$. Nous avons $\forall\, d \in I\!R^n$, $\forall\, \lambda > 0$
$f(x_0 + \lambda d) \geq f(x_0) + \lambda < x_0^*, d > + \frac{\lambda^2}{2} < \underline{\nabla}^2 f(x_0) d, d > + o(\lambda^2)$

Démonstration :

Nous savons, par construction même de la dérivée directionnelle du second ordre de f en x_0 [cfr (1)], que

$$f(x_0 + \lambda d) = f(x_0) + \lambda f'(x_0; d) + \frac{\lambda^2}{2} f''(x_0; d) + o(\lambda^2)$$

En vertu de la proposition A.IV.3, nous déduisons

$$f(x_0 + \lambda d) \geq f(x_0) + \lambda < x_0^*, d > + \frac{\lambda^2}{2} f''(x_0; d) + o(\lambda^2)$$

En outre, par définition,

$$\underline{E} = [\underline{\nabla}^2 f(x_0)]^{1/2}(\bar{B}) \subseteq \partial^2 f(x_0)$$

c'est-à-dire [cfr (31) et (32), ainsi que la proposition II.2.1]

$$\{x \mid \ < x, d > \leq \sqrt{< \underline{\nabla}^2 f(x_0) d, d >} \ \ \forall\, d\}$$

$$\subseteq \{x \mid \ < x, d > \leq \sqrt{f''(x_0; d)} \ \ \forall\, d\}$$

ce qui implique

$$< \underline{\nabla}^2 f(x_0) d, d > \leq f''(x_0; d) \ \ \forall\, d$$

La thèse en découle. ∎

CHAPITRE V

RELATION ENTRE LE DIFFERENTIEL SECOND ET LA MATRICE HESSIENNE GENERALISEE D'UNE FONCTION CONVEXE $C^{1,1}$

1 Définitions et résultats préliminaires

Rappelons qu'on désigne par fonction $C^{1,1}$, une fonction différentiable dont l'application gradient, ∇f, est localement lipschitzienne. Considérant une telle fonction, l'on définit, au sens de Clarke, la matrice hessienne de f au point $x_0 \in I\!R^n$ par
$Co\{M \mid \exists x_i \to x_0$ tel que f est deux fois différentiable en x_i et $\nabla^2 f(x_i) \to M\}$.
Cet ensemble est noté $\partial_c^2 f(x_0)$.

Proposition 1.1 :

$\partial_c^2 f(x_0)$ est un ensemble convexe, compact, non vide, constitué de matrices symétriques. Lorsque f est deux fois continûment différentiable, il se réduit à $\{\nabla^2 f(x_0)\}$.

On peut également concevoir une généralisation de la dérivée directionnelle de f au point x_0, soit $f^{00}(x_0; d)$, définie par

$$f^{oo}(x_0; d) = \lim_{\substack{x \to x_0 \\ \lambda \to 0^+}} \sup \frac{< \nabla f(x + \lambda d), d > \ - \ < \nabla f(x), d >}{\lambda}$$

Celle-ci est liée à la matrice hessienne généralisée au moyen de la propriété suivante.

Proposition 1.2 :

$$\underset{M \in \partial_c^2 f(x_0)}{\mathrm{Max}} < Md, d > = f^{00}(x_0; d) \ \forall \ d$$

(pour la preuve des deux affirmations précédentes nous nous référons à [9]).

Nous supposerons pour la suite que f est $C^{1,1}$ et convexe. Dès lors, les éléments de $\partial_c^2 f(x_0)$ deviennent semi-définis positifs (puisque limites de telles matrices) et nous pouvons associer à chacun d'eux un ellipsoïde $E = M^{1/2}(\bar{B})$ dont la fonction d'appui est $\sqrt{< Md, d >}$ [cfr chapitre 4, paragraphe 1].

2 Comparaison entre $\partial_c^2 f(x_0)$ et $\partial^2 f(x_0)$

Nous établissons tout d'abord deux propriétés desquelles nous pourrons déduire la relation souhaitée. La première compare entre elles les dérivées directionnelles du second ordre $f''(x_0; d)$ et $f^{00}(x_0; d)$. La seconde fait de $\sqrt{f^{00}(x_0; d)}$ une fonction d'appui.

Proposition 2.1 :

$$\sqrt{f^{00}(x_0; d)} \geq \sqrt{f''(x_0; d)} = \delta^*(\partial^2 f(x_0)|d) \ \forall \ d$$

Démonstration :

Nous savons en vertu de (2) que

$$\begin{aligned} f''(x_0; d) &= \lim_{\lambda \to 0^+} \frac{f'(x_0 + \lambda d; d) - f'(x_0; d)}{\lambda} \\ &= \lim_{\lambda \to 0^+} \frac{< \nabla f(x_0 + \lambda d), d > \ - \ < \nabla f(x_0), d >}{\lambda} \end{aligned}$$

lorsque f est différentiable [voir proposition A.IV.4].
Dès lors,

$$\begin{aligned} f''(x_0; d) &\leq \limsup_{\substack{x \to x_0 \\ \lambda \to 0^+}} \frac{< \nabla f(x + \lambda d) - \nabla f(x), d >}{\lambda} \\ &= f^{00}(x_0; d) \end{aligned}$$

par définition.
D'autre part, il résulte de la propriété A.III.4 que

$$\sqrt{f''(x_0; d)} = \delta^*(\partial^2 f(x_0) \mid d)$$

puisque $\sqrt{f''(x_0; d)}$ est convexe (f est différentiable), propre et semi-continue inférieurement. ■

Proposition 2.2 :

$$\sqrt{f^{00}(x_0; d)} = \sup_{z \in Co\left\{ \bigcup_{M \in \partial_c^2 f(x_0)} M^{1/2}(\bar{B}) \right\}} < z, d > \ \forall\, d$$

Démonstration :

Grâce à la proposition 1.2, il nous suffit de montrer que

$$\sqrt{\max_{M \in \partial_c^2 f(x_0)} < Md, d >} = \sup_{z \in Co(P)} < z, d >$$

où

$$P \overset{\text{not}}{=} \bigcup_{M \in \partial_c^2 f(x_0)} M^{1/2}(\bar{B})$$

a) $\forall\, M \in \partial_c^2 f(x_0)$, nous avons

$$\begin{aligned} \sqrt{< Md, d >} &= \sqrt{< M^{1/2}d, M^{1/2}d >} = ||M^{1/2}d|| \\ &= < M^{1/2}\frac{M^{1/2}d}{||M^{1/2}d||}, d > \end{aligned}$$

avec

$$M^{1/2}\frac{M^{1/2}d}{||M^{1/2}d||} \in M^{1/2}(\bar{B})$$

Ainsi,

$$\operatorname*{Max}_{M \in \partial_c^2 f(x_0)} \sqrt{< Md, d >} \leq \sup_{z \in Co(P)} < z, d >$$

b) Considérons un élément z arbitraire dans $Co(P)$. Il peut donc se représenter comme

$$z = \sum_{i=1}^{p} \lambda_i (M_i^{1/2}) b_i \qquad \lambda_i \geq 0, \sum_{i=1}^{p} \lambda_i = 1$$

où $b_i \in \bar{B}$ et $M_i \in \partial_c^2 f(x_0) \qquad \forall\, i = 1, \cdots, p.$

Dès lors,

$$\begin{aligned}
< z, d > &= < \sum_{i=1}^{p} \lambda_i M_i^{1/2} b_i, d > \\
&= \sum_{i=1}^{p} \lambda_i \ < M_i^{1/2} b_i, d > \\
&= \sum_{i=1}^{p} \lambda_i < b_i, M_i^{1/2} d > \\
&\leq \sum_{i=1}^{p} \lambda_i \ ||b_i|| \ \ ||M_i^{1/2} d||
\end{aligned}$$

par l'égalité de Cauchy-Schwarz.

Comme $b_i \in \bar{B}$, nous avons encore

$$\begin{aligned}
< z, d > &\leq \sum_{i=1}^{p} \lambda_i \, ||M_i^{1/2} d|| \\
&= \sum_{i=1}^{p} \lambda_i \sqrt{< M_i d, d >} \\
&\leq \max_{M \in \partial_c^2 f(x_0)} \sqrt{< M d, d >} \sum_{i=1}^{p} \lambda_i \\
&= \max_{M \in \partial_c^2 f(x_0)} \sqrt{< M d, d >}
\end{aligned}$$

Par conséquent,

$$\sup_{z \in Co(P)} < z, d > \leq \operatorname*{Max}_{M \in \partial_c^2 f(x_0)} \sqrt{< M d, d >}$$

∎

Rassemblant les résultats des propositions 2.1 et 2.2, nous obtenons que

$$\delta^*(\partial^2 f(x_0) \mid d) \leq \delta^*(Co\{ \bigcup_{M \in \partial_c^2 f(x_0)} M^{1/2}(\bar{B})\} | d) \ \ \forall \ d$$

Nous nous référons alors à la proposition A.VIII.1 pour déduire que $\partial^2 f(x_0)$ est contenu dans l'enveloppe convexe des ellipsoïdes associés aux matrices de $\partial_c^2 f(x_0)$ c'est-à-dire

$$\partial^2 f(x_0) \subseteq Co\{ \bigcup_{M \in \partial_c^2 f(x_0)} M^{1/2}(\bar{B})\}$$

De plus, comme le membre de droite de cette inclusion est symétrique par rapport à l'origine, nous avons en fait que

$$Co\{\partial^2 f(x_0) \cup -\partial^2 f(x_0)\} \subseteq Co\{\bigcup_{M\in\partial_c^2 f(x_0)} M^{1/2}(\bar{B})\} \tag{33}$$

Nous considérons à présent un cas particulier de fonctions convexes $C^{1,1}$ pour lequel l'égalité se produit dans la relation (33).
Il s'agit de la fonction $f : I\!R^n \to I\!R$ définie par

$$f(x) = \frac{1}{2}[g^+(x)]^2$$

où g est une fonction deux fois continûment différentiable et convexe. Par la notation $g^+(x)$, nous entendons $\max(0, g(x))$.
Nous supposons que

$$\exists\ \bar{x} \text{ tel que } g(\bar{x}) < 0$$

sans quoi $f(x) = \frac{1}{2}[g(x)]^2$ constituerait une fonction deux fois continûment différentiable et l'égalité serait trivialement vérifiée [cfr proposition 1.1 et II.2.5].
Prenons un point $x_0 \in I\!R^n$ tel que $g(x_0) = 0$. Nous obtenons alors que le gradient $\nabla g(x_0) \neq 0$ puisque x_0 n'est pas minimum de g.
Nous avons que

$$\nabla f(x) = \nabla g(x)\ g^+(x) \quad \text{et} \quad \nabla f(x_0) = 0 \tag{34}$$

En effet,

a) Si $g(x) > 0$, alors

$$f(x) = \frac{1}{2}[g(x)]^2$$

et

$$\nabla f(x) = g(x)\nabla g(x)$$

b) Si $g(x) < 0$, alors

$$f(x) = 0$$

et

$$\nabla f(x) = 0$$

c) Si $g(x) = 0$, alors nous devons vérifier que pour $\nabla f(x) = \nabla g(x)\ g^+(x)$,

$$< \nabla f(x), d >= \lim_{\lambda \to 0} \frac{f(x + \lambda d) - f(x)}{\lambda}$$

Or,

$$\begin{aligned}
\lim_{\lambda\to 0}\frac{f(x+\lambda d)-f(x)}{\lambda} &= \lim_{\lambda\to 0}\frac{\frac{1}{2}([g^+(x+\lambda d)]^2-[g^+(x)]^2)}{\lambda}\\
&= \lim_{\lambda\to 0}\frac{1}{2}\frac{([\max(0,g(x+\lambda d))]^2-0)}{\lambda}
\end{aligned}$$

Envisageons donc l'alternative suivante.

(i) Si $g(x+\lambda d)>0$:

$$\begin{aligned}
\lim_{\lambda\to 0}\frac{1}{2}\frac{[\max(0,g(x+\lambda d))]^2}{\lambda} &= \lim_{\lambda\to 0}\frac{1}{2}\frac{g^2(x+\lambda d)}{\lambda}\\
&= 0
\end{aligned}$$

puisque g est continue.

(ii) Si $g(x+\lambda d)\leq 0$:

$$\lim_{\lambda\to 0}\frac{1}{2}\frac{[\max(0,g(x+\lambda d)]^2}{\lambda}=0$$

Dès lors, nous obtenons

$$\begin{aligned}
<\nabla f(x),d> &= g^+(x)\ <\nabla g(x),d>\\
&= 0\\
&= \lim_{\lambda\to 0}\frac{f(x+\lambda d)-f(x)}{\lambda}
\end{aligned}$$

et le résultat (34) est démontré. □

Ainsi, pour établir le caractère $C^{1,1}$ de la fonction f, il ne nous reste plus qu'à prouver la propriété locale de Lipschitz pour l'application gradient ∇f. Il nous faut donc montrer que

$$\begin{aligned}
&\exists\varepsilon>0,\exists K>0 \text{ tel que } \forall\ x,y\in\bar{B}(x_1,\varepsilon)\\
&\qquad ||\nabla f(y)-\nabla f(x)||\leq K\ ||y-x||
\end{aligned}\tag{35}$$

où x_1 est un point arbitraire de $I\!R^n$.
Pour ce faire, envisageons trois cas.

1. $g(x_1)>0$:

Sous cette hypothèse, nous pouvons affirmer que

$$\exists\varepsilon_1>0 \text{ tel que } \forall\ x'\in\bar{B}(x_1,\varepsilon_1)\quad g(x')>0$$

Dès lors, $g^+(x') = g(x')$ et $\forall\ x, y \in \bar{B}(x_1, \varepsilon)$, nous avons

$$\begin{aligned} \|\nabla f(y) - \nabla f(x)\| &= \|g(y)\nabla g(y) - g(x)\nabla g(x)\| \\ &\leq |g(y)|\|\nabla g(y) - \nabla g(x)\| + |g(y) - g(x)|\ \|\nabla g(x)\| \\ &\leq M_1 K_2\|y - x\| + K_3\|y - x\|M_2 \end{aligned}$$

car g est deux fois continûment différentiable, ce qui implique

(a) ∇g localement lipschitzien en x_1 :
Nous avons ainsi que

$$\begin{aligned} &\exists \varepsilon_2 > 0\ \exists K_2 > 0 \text{ tel que } \forall\ x, y \in \bar{B}(x_1, \varepsilon_2) \\ &\|\nabla g(y) - \nabla g(x)\|\ \leq K_2\ \|y - x\| \end{aligned}$$

(b) g continue sur $\bar{B}(x_1, \varepsilon_1)$:
Donc, g est bornée c'est-à-dire

$$\exists\ M_1 > 0 \text{ tel que } \forall\ y \in \bar{B}(x_1, \varepsilon_1)\quad |g(y)| \leq M_1$$

(c) g localement lipschitzienne en x_1 :
D'où, $\exists \varepsilon_3 > 0\ \exists K_3 > 0$ tel que $\forall\ x, y \in \bar{B}(x_1, \varepsilon_3)$

$$|g(y) - g(x)|\ \leq K_3\ \|y - x\|$$

(d) ∇g continue sur $\bar{B}(x_1, \varepsilon_1)$:
Par conséquent, ∇g est bornée. Donc,

$$\exists M_2 > 0 \text{ tel que } \forall\ x \in \bar{B}(x_1, \varepsilon_1)\quad \|\nabla g(x)\|\ \leq M_2$$

Finalement, nous obtenons (35) avec

$$\begin{aligned} & K = M_1 K_2 + M_2 K_3 \\ \text{et}\quad & \varepsilon = \min \varepsilon_i \quad i = 1, 2, 3 \end{aligned}$$

2. $g(x_1) < 0$:
Cette fois, $\exists \varepsilon > 0$ tel que $\forall\ x' \in \bar{B}(x_1, \varepsilon)\ \ g(x') < 0$.
Consécutivement, $g^+(x') = 0$ et $\forall\ x, y \in \bar{B}(x_1, \varepsilon)$, nous avons

$$\begin{aligned} \|\nabla f(y) - \nabla f(x)\| &= 0 \\ &\leq K\|y - x\| \quad \forall\ K > 0 \end{aligned}$$

3. $g(x_1) = 0$:

Dans ce cas, $\exists \varepsilon > 0$ tel que pour x et y choisis arbitrairement dans $\bar{B}(x_1, \varepsilon)$, nous avons

(a) $g(y) \leq 0$ et $g(x) \leq 0$: cette possibilité a déjà été étudiée en 2.

ou bien

(b) $g(y) > 0$ et $g(x) > 0$: cette possibilité correspond au cas 1.

ou bien encore

(c) $g(y) > 0$ et $g(x) \leq 0$:

(Le cas $g(y) \leq 0$ et $g(x) > 0$ est similaire).

Sous cette hypothèse, nous obtenons, au moyen du théorème de la valeur intermédiaire que

$$\begin{aligned} ||\nabla f(y) - \nabla f(x)|| &= ||\nabla g(y) g(y) - 0|| \\ &= ||\nabla g(y) g(y) - \nabla g(y) g(z)|| \end{aligned}$$

où $z \in [x, y]$ est tel que $g(z) = 0$.

Ainsi,

$$\begin{aligned} ||\nabla f(y) - \nabla f(x)|| &= ||\nabla g(y)||\ |g(y) - g(z)| \\ &\leq M_5 K_5\ ||y - z|| \end{aligned}$$

puisque ∇g est continue sur $\bar{B}(x_1, \varepsilon)$ (donc bornée) et g est localement lipschitzienne en x_1.

Comme $z \in [x, y]$, nous avons encore que

$$||\nabla f(y) - \nabla f(x)|| \leq M_5 K_5\ ||y - x||$$

ce qui nous conduit à (35). □

Montrons maintenant le caractère convexe de la fonction f au moyen des deux étapes ci-dessous.

a) Si g est convexe, alors g^+ l'est également :

En effet, nous désirons établir que $\forall\ \lambda \in [0, 1]$

$$g^+(\lambda x + (1 - \lambda) y) \leq \lambda g^+(x) + (1 - \lambda) g^+(y)$$

Nous considérons dès lors les deux cas suivants.

1^0) Si $\lambda x + (1-\lambda)y \geq 0$:

Sous cette hypothèse, nous avons

$$\begin{aligned} g^+(\lambda x + (1-\lambda)y) &= g(\lambda x + (1-\lambda)y) \\ &\leq \lambda g(x) + (1-\lambda)g(y) \\ &\leq \lambda g^+(x) + (1-\lambda)g^+(y) \end{aligned}$$

car g est convexe et $g \leq g^+$.

2^0) Si $\lambda x + (1-\lambda)y < 0$:

Dans ce cas, nous avons

$$0 \leq \lambda g^+(x) + (1-\lambda)g^+(y)$$

puisque chacun des termes du membre de droite est positif.

b) Si g^+ est convexe, alors $(g^+)^2$ le devient aussi.

De fait, par la convexité de g^+, nous obtenons

$$\begin{aligned} [g^+(\lambda x + (1-\lambda)y)]^2 &\leq [\lambda g^+(x) + (1-\lambda)g^+(y)]^2 \\ &= \lambda^2[g^+(x)]^2 + (1-\lambda)^2[g^+(y)]^2 + 2\lambda(1-\lambda)g^+(x)g^+(y) \end{aligned}$$

Pour obtenir le résultat annoncé, nous devons donc prouver que

$\lambda^2[g^+(x)]^2 + (1-\lambda)^2[g^+(y)]^2 + 2\lambda(1-\lambda)g^+(x)g^+(y)$

$$-\lambda[g^+(x)]^2 - (1-\lambda)[g^+(y)]^2 \leq 0$$

c'est-à-dire

$\lambda(\lambda-1)[g^+(x)]^2 + [1 - 2\lambda + \lambda^2 - 1 + \lambda][g^+(y)]^2$

$$+2\lambda(1-\lambda)g^+(x)g^+(y) \leq 0$$

ou encore

$$\lambda(\lambda-1)([g^+(x)]^2 + [g^+(y)]^2 + 2g^+(x)g^+(y)) \leq 0$$

Comme

$$[g^+(x)]^2 + [g^+(y)]^2 + 2g^+(x)g^+(y) = [g^+(x) + g^+(y)]^2 \geq 0$$

cette inégalité est bien satisfaite. □

Ayant ainsi vérifié les caractères $C^{1,1}$ et convexe de la fonction f, nous sommes en mesure d'en calculer la matrice hesienne généralisée, au sens de Clarke, au point x_0.
Nous savons, par définition, que
$\partial_c^2 f(x_0) = Co\{M \mid \exists x_i \to x_0$ tel que f est deux fois différentiable en x_i et $\nabla^2 f(x_i) \to M\}$
Envisageons les deux possibilités suivantes.

a) Si $g(x_i) > 0$:
Alors $f(x_i) = \frac{1}{2}[g(x_i)]^2$ est deux fois différentiable et comme $\nabla f(x_i) = g(x_i)\nabla g(x_i)$, nous obtenons

$$\nabla^2 f(x_i) = g(x_i)\nabla^2 g(x_i) + \nabla g(x_i)\nabla g(x_i)^T$$

b) Si $g(x_i) < 0$:
Alors $f(x_i) = 0$ et $\nabla^2 f(x_i) = 0$.
Consécutivement, si $x_i \to x_0$, alors, comme g est deux fois continûment différentiable, nous obtenons
$\nabla^2 f(x_i) \to g(x_0)\nabla^2 g(x_0) + \nabla g(x_0)\nabla g(x_0)^T$ cas a)
ou
$\nabla^2 f(x_i) \to 0$ cas b)
Puisque $g(x_0) = 0$, nous avons finalement que

$$\begin{aligned}\partial_c^2 f(x_0) &= Co\{0, \nabla g(x_0)\nabla g(x_0)^T\}\\ &= \{\alpha\nabla g(x_0)\nabla g(x_0)^T \mid \alpha \in [0,1]\}\end{aligned}$$

□

Il suit de ce résultat que

$$f^{00}(x_0; d) = (<\nabla g(x_0), d>)^2$$

En effet, de la proposition 1.2, nous déduisons

$$\begin{aligned}f^{00}(x_0; d) &= \max_{M\in\{\alpha\nabla g(x_0)\nabla g(x_0)^T \mid \alpha\in[0,1]\}} <Md, d>\\ &= \max_{\alpha\in[0,1]} <\alpha\nabla g(x_0)\nabla g(x_0)^T d, d>\\ &= <\nabla g(x_0)\nabla g(x_0)^T d, d> \max_{\alpha\in[0,1]} \alpha\\ &= <\nabla g(x_0)^T d, \nabla g(x_0)^T d>\\ &= << \nabla g(x_0), d> , <\nabla g(x_0), d>>\\ &= <\nabla g(x_0), d> \,.\, <\nabla g(x_0), d>\\ &= (<\nabla g(x_0), d>)^2\end{aligned}$$

□

Nous calculons à présent la dérivée directionnelle du second ordre de f en x_0 dans la direction d.
Dans ce but, nous considérons l'alternative suivante.

a) Si $\nabla g(x_0)^T d < 0$:
Alors, comme f est différentiable, nous obtenons

$$\begin{aligned} f''(x_0; d) &= \lim_{\lambda \to 0^+} \frac{\nabla f(x_0 + \lambda d)^T d - \nabla f(x_0)^T d}{\lambda} \\ &= \lim_{\lambda \to 0^+} \frac{\nabla f(x_0 + \lambda d)^T d}{\lambda} && \text{[cfr (34)]} \\ &= 0 \end{aligned}$$

car

$$\begin{aligned} \nabla f(x_0 + \lambda d)^T d &= g^+(x_0 + \lambda d)\ \ \nabla g(x_0 + \lambda d)^T d \\ &= 0 \end{aligned}$$

pour λ suffisamment petit.

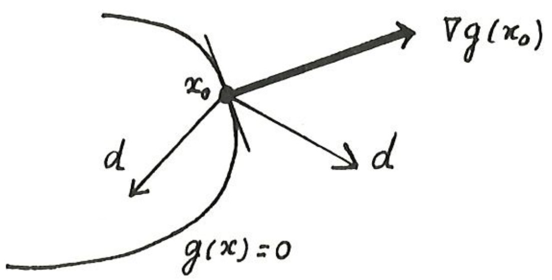


Figure 26

b) Si $\nabla g(x_0)^T d \geq 0$:
Dans ce cas,

$$\begin{aligned} f''(x_0; d) &= \lim_{\lambda \to 0^+} \frac{\nabla f(x_0 + \lambda d)^T d - 0}{\lambda} \\ &= \lim_{\lambda \to 0^+} \frac{\nabla g(x_0 + \lambda d)^T d\ g(x_0 + \lambda d)}{\lambda} \\ &= \lim_{\lambda \to 0^+} \nabla g(x_0 + \lambda d)^T d \cdot \lim_{\lambda \to 0^+} \frac{g(x_0 + \lambda d) - g(x_0)}{\lambda} \end{aligned}$$

parce que $g(x_0) = 0$.

Dès lors, puisque g est deux fois continûment différentiable, nous avons

$$\begin{aligned} f''(x_0; d) &= [\nabla g(x_0)^T d][\nabla g(x_0)^T d] \\ &= (< \nabla g(x_0), d >)^2 \end{aligned}$$

Rassemblant les résultats obtenus, nous pouvons dire que

$$f''(x_0; d) = [< \nabla g(x_0), d >^+]^2 \tag{36}$$

□

Montrons finalement que

$$\partial^2 f(x_0) = Co\{0, \nabla g(x_0)\}$$

Nous savons, par la propriété A.VIII.1, qu'il nous suffit de prouver l'égalité des fonctions d'appui.

Or,

$$\begin{aligned} \sup_{x^{**} \in Co\{0, \nabla g(x_0)\}} < x^{**}, d > &= \sup_{0 \leq \alpha \leq 1} < \alpha \nabla g(x_0), d > \\ &= [< \nabla g(x_0), d >]^+ \\ &= \sqrt{f''(x_0; d)} \qquad \text{[cfr (36)]} \end{aligned}$$

Dès lors, en faisant appel à la proposition 2.1, nous obtenons l'égalité souhaitée. □

Observons maintenant que l'ellipsoïde E_α associé à $M_\alpha = \alpha^2 \nabla g(x_0) \nabla g(x_0)^T \in \partial_c^2 f(x_0)$ n'est autre que

$$Co\{-\alpha \nabla g(x_0), \alpha \nabla g(x_0)\} \tag{37}$$

En effet, pour simplifier l'écriture nous notons

$$u = \alpha \nabla g(x_0)$$

où, sans perte de généralité, nous pouvons supposer $\alpha \neq 0$.

Nous devons donc démontrer que

$$Co\{-u, u\} = [uu^T]^{1/2}(\bar{B})$$

Pour ce faire, nous considérons la matrice R de similitude (composée d'une rotation de centre 0 et d'une homothétie) telle que

$$Ru = e_1$$

où e_1 désigne le premier vecteur de la base canonique de $I\!R^n$.

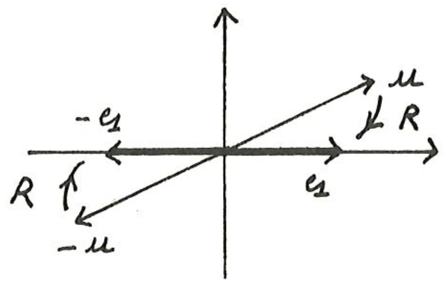


Figure 27

Comme R est inversible, il nous suffit de prouver que

$$R(uu^T)^{1/2}\overline{B} = R(Co\,\{-u, u\})$$

Nous avons, d'une part, que

$$R(Co\,\{-u, u\}) = Co\,\{-e_1, e_1\}$$

et, d'autre part, que

$$R(uu^T)^{1/2} = (Ru(Ru)^T)^{1/2}$$

En effet, par l'unicité de la racine carrée d'une matrice semi-définie positive, il est suffisant de montrer que les carrés coïncident, c'est-à-dire

$$[R(uu^T)^{1/2}]^2 = Ru(Ru)^T$$

ce qui revient à l'égalité triviale suivante :

$$R(uu^T)^{1/2}(uu^T)^{1/2}R^T = Ruu^T R^T$$

Ainsi, il nous reste à avoir que

$$(Ru(Ru)^T)^{1/2}\overline{B} = Co\,\{-e_1, e_1\}$$

Ceci s'obtient facilement en constatant que

$$\begin{aligned}(Ru(Ru)^T)^{1/2}\overline{B} &= (e_1 e_1^T)^{1/2}\overline{B} \\ &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}\overline{B}\end{aligned}$$

$\square$

Consécutivement, nous obtenons que

$$Co\{\bigcup_{M\in\partial_c^2 f(x_0)} M^{1/2}(\bar{B}\} = Co(-\nabla g(x_0), \nabla g(x_0))$$

En effet, la première inclusion ($\supseteq$) est immédiate. Nous démontrons la seconde. Considérons tout d'abord un élément de la forme

$$z = M^{1/2}b \text{ avec } \begin{cases} ||b|| \leq 1 \\ M \in \partial_c^2 f(x_0) \end{cases}$$

Nous avons que

$$\begin{aligned} z = M^{1/2}b &= \alpha\nabla g(x_0)\nabla g(x_0)^T b \quad \text{avec } \alpha \in [0,1] \\ &\leq \nabla g(x_0)\nabla g(x_0)^T b \end{aligned}$$

Dès lors,

$$z \in (M_1)^{1/2}(\bar{B}) = Co(-\nabla g(x_0), \nabla g(x_0)) \tag{38}$$

en vertu de l'égalité (37).
Ainsi, pour

$$z \in Co\{\bigcup_{M\in\partial_c^2 f(x_0)} M^{1/2}(\bar{B})\}$$

c'est-à-dire, avec

$$\lambda_i \geq 0 \quad i = 1, \cdots, p \text{ et } \sum_{i=1}^{p} \lambda_i = 1$$

$$z = \sum_{i=1}^{p} \lambda_i (M_i)^{1/2} b_i = \sum_{i=1}^{p} \lambda_i z_i$$

où

$$z_i \in (M_i)^{1/2}(\bar{B}) \subseteq (M_1)^{1/2}(\bar{B}) = Co(-\nabla g(x_0), \nabla g(x_0)) \qquad \text{[cfr (38)]}$$

Finalement, nous obtenons, par définition de l'enveloppe convexe,

$$\sum_{i=1}^{p} \lambda_i z_i \in Co(-\nabla g(x_0), \nabla g(x_0))$$

$\square$

D'autre part, puisque

$$\partial^2 f(x_0) = Co\{0, \nabla g(x_0)\}$$

nous déduisons que

$$Co(\partial^2 f(x_0) \cup -\partial^2 f(x_0)) = Co\{-\nabla g(x_0), \nabla g(x_0)\}$$

Ainsi, nous constatons que nous avons pu établir l'égalité dans (33). $\square$

Nous remarquons, de plus, que les dérivées secondes intérieures et extérieures sont les suivantes

$$\begin{aligned} & \underline{\nabla}^2 f(x_0) = \{0\} \\ \text{et} \quad & \overline{\nabla}^2 f(x_0) = \nabla g(x_0) \nabla g(x_0)^T \end{aligned}$$

$\blacksquare$

ANNEXE

Nous reprenons ci-dessous quelques définitions et propriétés fondamentales de l'analyse convexe du premier ordre.

I. Ensembles et fonctions convexes

I.1. Un *ensemble* $C \subset \mathbb{R}^n$ est *convexe* si $\forall\, x, y \in C \quad \forall\, \lambda \in [0,1]$, on a

$$\lambda x + (1-\lambda) y \in C$$

I.2. *Proposition* : Si $\mathcal{F}$ est une famille d'ensembles convexes de $\mathbb{R}^n$. Alors $\bigcap_{C \in \mathcal{F}} C$ est un ensemble convexe.

I.3. Soient un ensemble $S \subseteq \mathbb{R}^n$ et $\mathcal{F}$, la famille de toutes les parties convexes dans $\mathbb{R}^n$ contenant S ($\mathcal{F}$ est non vide car $\mathbb{R}^n$ y est inclus). On considère $Co(S) \equiv \bigcap_{C \in \mathcal{F}} C$ qu'on appelle *enveloppe convexe* de S. Il s'agit du plus petit ensemble convexe contenant S.

Proposition : Supposons S non vide. Nous avons que

$$x \in Co(S) \Leftrightarrow \exists x_i \in S, \exists \lambda_i \geq 0 \quad i = 1, \cdots, k \text{ avec } \sum_{i=1}^{k} \lambda_i = 1$$

$$\text{tels que} \qquad x = \sum_{i=1}^{k} \lambda_i x_i$$

I.4. Soit $C \subset \mathbb{R}^n$, un ensemble convexe non vide dans lequel on considère un point x. Ce point est dit *point extrémal* de C si
$\forall\, y, z \in C \quad \forall\, \lambda \in [0,1]$

$$x = \lambda y + (1-\lambda) z \Rightarrow x = y = z$$

Considérons dorénavant la fonction $f : \mathbb{R}^n \to \overline{\mathbb{R}}$.

I.5. On appelle *épigraphe* de f, noté *epi* (f), l'ensemble

$$\{(x,a) \subseteq \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq a\}$$

I.6. On définit le *domaine* de f par

$$\text{dom } f = \{x \in I\!R^n \mid f(x) < +\infty\}$$

I.7. f est dite *propre* si et seulement si

$$\text{dom } f \neq \emptyset \text{ et } f(x) > -\infty \quad \forall\, x \in I\!R^n$$

I.8. f est une *fonction convexe* si et seulement si *epi* (f) est un ensemble convexe de $I\!R^{n+1}$.

I.9. *Proposition* : Si f est propre, alors

$$f \text{ convexe } \Leftrightarrow \begin{cases} \text{dom } f \text{ est convexe et} \\ \forall\, x, y \in \text{dom } f \quad \forall\, \lambda \in [0,1] \\ f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \end{cases}$$

I.10. *Proposition* : Soit $\mathcal{F}$ une famille de fonctions convexes sur $I\!R^n$.
Considérons $f(x) \equiv \sup\limits_{g \in \mathcal{F}} g(x) \quad \forall\, x \in I\!R^n$.
Alors f est convexe.

I.11. *L'intérieur relatif* d'un ensemble est l'intérieur de cet ensemble relativement à la variété affine engendrée par cet ensemble.

I.12. *Proposition* : Toute fonction convexe est continue sur l'intérieur relatif de son domaine.

I.13. *Corollaire* : Toute fonction convexe de domaine $I\!R^n$ à valeurs dans $I\!R$ est continue.

II. Semi-continuité inférieure et fermeture

II.1. La fonction $f : I\!R^n \to \overline{I\!R}$ est *semi-continue inférieurement* en $x_0 \in I\!R^n$ si

$$\forall\, \lambda < f(x_0)\ \exists \delta > 0 \text{ tel que } \forall\, x \in I\!R^n$$

$$||x - x_0|| < \delta \Rightarrow \lambda < f(x)$$

II.2. *Proposition* : Soit $\mathcal{F}$ une famille de fonctions semi-continues inférieurement sur $I\!R^n$.
Considérons

$$f(x) = \sup_{g \in \mathcal{F}} g(x) \quad \forall\, x \in I\!R^n$$

Alors f est semi-continue inférieurement.

II.3. On définit *la fermeture* de la fonction convexe propre $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ par $cl : \mathbb{R}^n \to \overline{\mathbb{R}}$ telle que

$$(cl\ f)(x) = \sup_{\substack{F \text{ affine} \\ F \leq f}} F(x)$$

Nous remplacerons dorénavant $(cl\ f)(x)$ par $cl\ f(x)$ (abus de notation).

II.4. *Proposition* : $cl\ f$ est semi-continue inférieurement et

$$epi\ (cl\ f) = \overline{epi\ (f)}$$

II.5. *Proposition* : Soit $f : \mathbb{R}^n \to \mathbb{R}$ convexe propre.
Alors on a l'équivalence entre les propriétés suivantes

(i) f est semi-continue inférieurement

(ii) $epi\ f$ est fermé dans $\mathbb{R}^{n+1}$

(iii) $\forall\ a \in \mathbb{R}$, l'ensemble niveau $\{x \mid f(x) \leq a\}$ est fermé dans $\mathbb{R}^n$.

(iv) $cl\ f = f$ c'est-à-dire f est *fermée*

II.6. *Proposition* : $cl\ f$ est convexe et $cl\ f(x) \leq f(x) \quad \forall\ x$

II.7. *Proposition* : Soit à nouveau $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexe, propre. Nous avons que

$$(cl\ f)(x) = \lim_{y \to x} \inf f(y)$$

II.8. *Proposition* : La fonction $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ est semi-continue inférieurement si et seulement si

$$f(x) = \lim_{y \to x} \inf f(y)$$

La continuité engendre donc bien la semi-continuité inférieure.

III. Fonctions conjuguées

Soit une fonction $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ propre.

III.1. On appelle *conjuguée de* f, la fonction de $\mathbb{R}^n$ dans $\overline{\mathbb{R}}$, notée f^*, définie par

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \{< y, x > - f(x)\}$$

III.2. *Proposition* : f^* est convexe, propre, semi-continue inférieurement.

III.3. *Proposition* : Si f est convexe, la biconjuguée de f, soit f^{**}, satisfait l'égalité

$$f^{**} = cl\ f$$

III.4. *Proposition* : Si la fonction f est convexe, propre et semi-continue inférieurement, alors $f^{**} = f$.

IV. propriétés de la dérivée directionnelle et du sous-différentiel

Considérons la fonction $f : I\!R^n \to I\!R$ convexe et un point x_0 de $I\!R^n$. Les définitions ont été rappelées dans le chapitre I paragraphe 1. Nous citons ici quelques propriétés intéressantes.

IV.1. *Proposition* : $\partial f(x_0)$ est un convexe fermé, borné, non vide.

IV.2. : $d \to f'(x_0; d)$ est convexe et positivement homogène de degré un (ce qui signifie : $\forall\ \alpha \geq 0\ \ f'(x_0; \alpha d) = \alpha f'(x_0; d)$).

IV.3. *Proposition* : x^* est un sous-gradient de f en x_0, c'est-à-dire $x^* \in \partial f(x_0)$ si et seulement si $f'(x_0; d) \geq < x^*, d > \quad \forall\ d \in I\!R^n$

IV.4. *Proposition* : Si f est différentiable, alors

$$f'(x_0; d) = < \nabla f(x_0), d > \quad \text{et} \quad \partial f(x_0) = \{\nabla f(x_0)\}$$

IV.5. *Proposition* : Nous avons que

$$f'(x_0; d) = \delta^*(\partial f(x_0) \mid d) = \sup_{x^* \in \partial f(x_0)} < x^*, d >$$

Pour la définition détaillée de la fonction d'appui, nous nous référons au chapitre I paragraphe 2.4.1.

IV.6. *Proposition* : [Calculs concernant le sous-différentiel]

a) Soient $f, g : I\!R^n \to I\!R \cup \{+\infty\}$ convexes. Alors

$$\partial(f+g)(x) \supseteq \partial f(x) + \partial g(x) \quad \forall\ x \in I\!R^n$$

Si, de plus, f est continue en un point, alors nous obtenons l'égalité pour tout x.

b) Soient les fonctions $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ convexes $i = 1, \cdots, p$ et un point x de $\mathbb{R}^n$. Si f_i est continue en x $i = 1, \cdots, p$ alors

$$\partial(\sup_{1 \leq i \leq p} f_i)(x) = Co \bigcup_{i \in I(x)} \partial f_i(x) \quad \forall\, x \in \mathbb{R}^n$$

où

$$I(x) = \{i = 1, \cdots, n \mid f_i(x) = \sup_{1 \leq j \leq p} f_j(x)\}$$

c) Si $h : \mathbb{R} \to \mathbb{R}$ est continûment différentiable, alors

$$\partial\,(h \circ f)(x) = h'(f(x))\partial f(x)$$

IV.7. *Proposition* : Si $f(x) = \max_{i=1,\cdots,p} f_i(x)$ où les $f_i : \mathbb{R}^n \to \mathbb{R}$ constituent des fonctions convexes, alors pour $x_0, d \in \mathbb{R}^n$, nous avons

$$f'(x_0; d) = \max_{i \in I_0} f_i'(x_0; d)$$

où

$$I(x_0) = \{i \mid f_i(x_0) = f(x_0)\}$$

IV.8. *Proposition* : Soit $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ une fonction convexe, propre, semi-continue inférieurement. Alors

$$x^* \in \partial f(x) \Leftrightarrow x \in \partial f^*(x^*)$$

V. Cônes normaux et tangents

V.1. Soit un convexe $K \subseteq \mathbb{R}^n$ et un point $x \in K$
Le *cône normal* à K en x, noté $N(K, x)$ est défini par

$$N(K, x) = \partial\delta(K \mid .)(x)$$

où $\delta(K \mid .)$ désigne la fonction indicatrice de K [voir chapitre I paragraphe 2.4.1]. Ainsi,

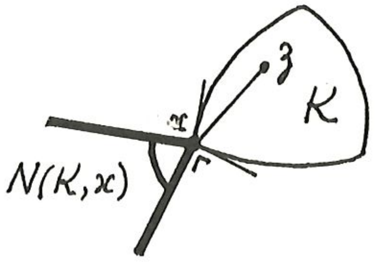
$$x^* \in \partial\delta(K \mid .)(x) \Leftrightarrow \;\; \forall\, z \in \mathbb{R}^n \;\; \delta(K \mid z) \geq \delta(K \mid x) + < x^*, z - x >$$

et, puisque $x \in K$

$$x^* \in \partial\delta(K \mid .)(x) \Leftrightarrow \;\; \forall\, z \in \mathbb{R}^n \;\; \delta(K \mid z) \geq 0 + < x^*, z - x >$$

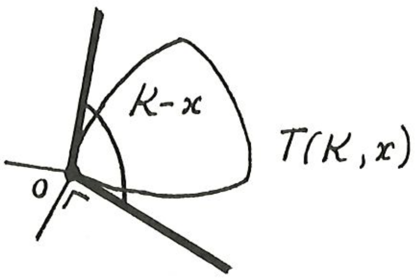
c'est-à-dire

$$x^* \in \partial\delta(K \mid .)(x) \Leftrightarrow \forall\, z \in K \;\; 0 \geq \, < x^*, z - x >$$



I.2. Le *cône tangent* à K en x *est défini par*

$$T(K,x) = \overline{\bigcup_{h>0} \frac{1}{h}(K - x)}$$



Nous avons $T(K,x)^- = N(K,x) \quad \forall\, x \in K$
où $A^- = \{x \in I\!R^n \mid \, < x, a > \leq 0 \quad \forall\, a \in A\}$.

VI. Propriétés extrémales des fonctions convexes

Considérons le problème suivant

$$\begin{cases} \text{Min} & f(x) \\ \text{sous} & x \in I\!R^n \end{cases}$$

lorsque $f : I\!R^n \to I\!R \cup \{+\infty\}$ est une fonction convexe.

VI.1. *Proposition* : Si f est propre et si x_0 est un point minimum local de f sur $I\!R^n$, alors x_0 est un point minimum global de f sur $I\!R^n$.

VI.2. *Proposition* : Soit $x_0 \in I\!R^n$. Nous avons que f atteint son minimem sur $I\!R^n$ en $x_0 \Leftrightarrow 0 \in \partial f(x_0)$.

VII. Les ellipsoïdes

Les définitions et propriétés fondamentales, liées à la géométrie des ellipsoïdes dans $I\!R^n$, ont été rappelées dans le chapitre IV paragraphe 1.

VIII. Compléments théoriques

VIII.1. *Proposition* : Soient A, B deux parties convexes fermées de $I\!R^n$

Si $\forall\, d \in I\!R^n$

$$\sup_{x^* \in A} < x^*, d > \leq \sup_{x^* \in B} < x^*, d >$$

Alors $A \subseteq B$.

REFERENCES

[1] BECKENBACH and R. BELLMAN, Inequalities, Springer Verlag, Berlin (1965), p. 63, theorem 6.

[2] H. BUSEMANN, Convex surfaces, Interscience Tracts in Pure and Applied Mathematics 6, Interscience Publishers, Inc., New York (1958), chapitre 1, paragraphe 2, (2.10) theorem.

[3] M. GRÖTSCHEL, L. LOVASZ and A. SCHRIJVER, Geometric algorithms and combinatorial optimization, Springer Verlag, Berlin (1988), p. 66.

[4] J.-B. HIRIART-URRUTY, *A new set-valued second order derivative for convex functions*, in J.-B. Hiriart Urruty, ed., Proceedings of the FERMAT days, North Holland (1986), p. 157-182.

[5] J.-B. HIRIAT-URRUTY, and A. SEEGER, *Calculus rules on a new set-valued second order derivative for convex functions*, preprint Laboratoire d'Analyse Numérique, Université Paul Sabatier, Toulouse (1985).

[6] J.-B. HIRIAT-URRUTY, J.-J. STRODIOT and V.H. NGUYEN, *Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data*, Applied Mathematics and Optimization, 11 (1984), p. 43-56.

[7] R.T. ROCKAFELLAR, Convex Analysis, princeton University Press (1970).

[8] A. SEEGER, *Analyse du second ordre de problèmes non-différentiables*, Thèse, Université Paul Sabatier (1986), chapitres B.1 et B.2.

[9] J. VAN TIEL, Convex Analysis : An Introductory Text, Wiley (1984).